

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: January 18, 2001, 06:46:44 ; Search time 14302.9 Seconds
(without alignments)
1479.199 Million cell updates/sec

Title: US-09-294-298-3
Perfect score: 4134
Sequence: 1 taagccccccaccgacc.....tgggtgtatcccatcct 4134

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 1118133 seqs, 2558875100 residues 2236266

Total number of hits satisfying chosen parameters: 2236266
Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba1.*
- 2: gb_ba2.*
- 3: gb_om.*
- 4: gb_ov.*
- 5: gb_ph.*
- 6: gb_pl1.*
- 7: gb_pl2.*
- 8: gb_pr1.*
- 9: gb_pr2.*
- 10: gb_pr3.*
- 11: em_fun.*
- 12: em_hum1.*
- 13: em_hum2.*
- 14: em_in.*
- 15: em_om.*
- 16: em_or.*
- 17: em_ov.*
- 18: em_pat.*
- 19: em_ph.*
- 20: em_pl.*
- 21: em_ro.*
- 22: em_sts.*
- 23: em_sy.*
- 24: em_un.*
- 25: em_v1.*
- 26: gb_htg1.*
- 27: gb_htg2.*
- 28: gb_in1.*
- 29: gb_in2.*
- 30: em_ba1.*
- 31: em_ba2.*
- 32: em_hum3.*
- 33: em_hum4.*
- 34: gb_pr4.*
- 35: gb_htg3.*
- 36: gb_htg4.*
- 37: gb_htg5.*
- 38: gb_htg6.*
- 39: gb_htg7.*
- 40: em_htg1.*
- 41: em_htg2.*
- 42: em_htg3.*
- 43: em_hum5.*

- 44: gb_pl3.*
- 45: gb_pr5.*
- 46: gb_htg8.*
- 47: gb_htg9.*
- 48: gb_htg10.*
- 49: gb_htg11.*
- 50: gb_htg12.*
- 51: gb_htg13.*
- 52: gb_htg14.*
- 53: gb_in3.*
- 54: gb_htg15.*
- 55: gb_htg16.*
- 56: gb_htg17.*
- 57: em_htg4.*
- 58: em_htg5.*
- 59: em_htg6.*
- 60: em_htg7.*
- 61: em_hum6.*
- 62: gb_htg18.*
- 63: gb_htg19.*
- 64: gb_ba3.*
- 65: em_htg8.*
- 66: em_htg9.*
- 67: em_htg10.*
- 68: gb_pr6.*
- 69: gb_pr7.*
- 70: gb_htg20.*
- 71: gb_htg21.*
- 72: gb_htg22.*
- 73: gb_htg23.*
- 74: gb_ro.*
- 75: gb_sts1.*
- 76: gb_sts2.*
- 77: gb_sy.*
- 78: gb_un.*
- 79: gb_v11.*
- 80: gb_v12.*
- 81: gb_pat1.*
- 82: gb_pat2.*
- 83: em_htg0.*
- 84: gb_htg24.*
- 85: gb_pr8.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %		DB ID	Description
		Match	Length		
1	4118.8	99.6	4140	74	AF058790 Rattus no
2	3919	94.8	4265	74	AF050183 Rattus no
3	3907.4	94.5	4539	74	AF058789 Rattus no
4	3680.4	89.0	4063	74	AF048976 Rattus no
5	3602.8	87.2	4801	74	AB016962 Rattus no
6	920.8	22.3	94770	69	AL050332 Human DNA
7	920.8	22.3	150956	70	AL161903 Homo sapi
8	920.8	22.3	169323	70	AL161906 Homo sapi
9	629	15.2	829	74	AF053938 Rattus no
10	627.8	15.2	4368	10	AF047711 Homo sapi
11	336.4	8.1	4287	34	AK024488 Homo sapi
12	241.8	5.8	56366	39	AC019800 Drosophil
13	241.8	5.8	191504	36	AC012162 Drosophil
14	241.8	5.8	202741	36	AC012161 Drosophil
15	241.8	5.8	300994	28	AE003506 Drosophil
16	236.8	5.7	3984	28	AB011280 Caenorhab
17	223.6	5.4	65961	38	AC016523 Drosophil
18	205.2	5.0	2385	34	AK022662 Homo sapi
19	196	4.7	172027	72	AL365274 Homo sapi
20	194.4	4.7	211	74	AF055883 Rattus no
21	187.6	4.5	169323	70	AL161906 Homo sapi

c 22 157 3.8 81971 68 HS593C16
23 157 3.8 156700 50 AC027051
24 139.2 3.4 155332 51 AC040900
c 25 139.2 3.4 156700 50 AC027051
26 113.4 2.7 4107 74 MM020238
27 103.4 2.5 159777 35 AC011703
28 103.2 2.5 157803 35 AC011492
29 100 2.4 39660 29 CEC07B5
30 99.2 2.4 2837 69 HSINSP4BP
31 98.4 2.4 237619 56 AC073765
32 97.6 2.4 833 28 AB011285
33 95.8 2.3 2647 81 EL1325
34 95.8 2.3 2730 31 BTU30857
35 93.8 2.3 760 28 AB011283
36 93.8 2.3 2276 28 AB011279
c 37 89.2 2.2 314 75 G21290
38 71.2 1.7 310 28 AB011284
c 39 70.8 1.7 5719 74 MWCATFS1
c 40 70.8 1.7 180385 9 AC007461
c 41 69.6 1.7 3720 80 S76368
c 42 69.6 1.7 43658 80 HSV3PRGEN
c 43 69.6 1.7 112930 80 HSGEND
c 44 68 1.6 5579 8 AB011110
c 45 67.4 1.6 314556 56 AC073495

ALIGNMENTS

RESULT 1
AF058790 AF058790 4140 bp mRNA ROD 09-OCT-1998
DEFINITION Rattus norvegicus SynGAP-b mRNA, complete cds.
ACCESSION AF058790
VERSION AF058790.1 GI:3722228
KEYWORDS Norway rat.
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 4140)
Kim, J.H., Liao, D., Lau, L.F. and Haganir, R.L.
SynGAP: a synaptic RASGAP that associates with the PSD-95/SAP90
protein family
Neuron 20 (4), 683-691 (1998)
98240917
2 (bases 1 to 4140)
Kim, J.H. and Haganir, R.L.
Direct Submission
Submitted (09-APR-1998) Neuroscience, Johns Hopkins Medical
Institute, 725 N. Wolfe Street, PCTB 900, Baltimore, MD 21205, USA
3 (bases 1 to 4140)
Kim, J.H. and Haganir, R.L.
Direct Submission
Submitted (08-OCT-1998) Neuroscience, Johns Hopkins Medical
Institute, 725 N. Wolfe Street, PCTB 900, Baltimore, MD 21205, USA
Sequence update by submitter
On Oct 9, 1998 this sequence version replaced gi:3065890.
Location/Qualifiers
1. .4140
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/tissue_type="hippocampus"
1. .3750
/note="neuronal RASGAP; contains PH domain, C2 domain, GAP
domain in the N-terminal region, and a T/SXV motif at the
C-terminus"
/codon_start=1
/product="SynGAP-b"
/protein_id="AAC63511.1"
/db_xref="GI:3722229"
/translation="MGLRPPTPTPPSGSGSLPPPSHRQPLRRRCSSCCFPGEYHLG

source

CDS

RSRKSPVGGKQYSMEAPAAPFRPSQGLSRRLKSSIKRTKYSQPKLDRTSSFRILP
RFRSDHRLRLMQSFKSHSHESLSFSSAAELMLNDEDSIIKPVHSSITLQGEFC
FEVTSSTCKFCFAAARAKRWENLQVAKPNKDNRRVNDLKLWLTIEARELPFK
KRYTCLELDDMLYARTTSKPRASNDFTVFWGEHFFNNLRPAVRLHLHYRDSKKR
KKGAGYGLVTPVATLAGRHFTQWYPTVLTPTGSGGSGMGSGGSGGSGGSGGK
KGGCPAVRLKARYOTMSILPMELYKEAEYVTHVYMLCAVLEPALNVKGEVASAL
VHILQSTCKAKDFLSDMAMSEVDREMERELHIFRENTLATNAEYMLRIGOKYUKDA
IGETRALYSESENEVDPIKCTASSLAHQANLRMCSELALCKVNVSHCVFPRELKE
VFASRLRCABEREDIDRLISASLFLRFLCPAITSFLGMEYFDEQTSRTTL
IAKIQNLANFSKFTSKEDFLGFNNEFLWMSQOFLYEISNLTDTNSSSPEGYI
DAGRELSTHALLMEVLPQLKGLPLRLSLDSTALNRDITNINQKQSRQSER
RSQMWLRGPSPAMOGYMWDRINSSDILQSFEMARGLNSSMDMARLPSPKPKPPPP
GGKDLFVVRPPLARSPACTSSSDITEPEOKMLSVNKYSVMDLDOGDGPGRLNS
SSVNLAAVGDLLHSSQASLTAALGLRPAPAGRLSQGSGSITAGMLRQMSQVTTDG
VPAQRLRPLSFQNLPHMAADGPPAGHSGSGSGGSGGSGGSGGSGGSGGSGGSGG
DTFAFHGYSKSEDLSTGVPKPPAAISILHSHSYDEFSGDTFRQLSLQDNLQHM
LSPPQITIGPQRPAPSGGSGGSGGSGGSGGSGGSGGSGGSGGSGGSGGSGGSGG
NLASPEPSYGPAPRQOGLSEKSGISGSGGSGGSGGSGGSGGSGGSGGSGGSGG
MPASERTVAVYNNMPLHSADIESAHIREEYKLKEYSKMDESLRDRVKEEETHSL
KERLHMSNRKLEEYERLLSQEOTSKILMOYQARLEQSEKRLRQOQVEKDSQIKSII
GRMLVEBELRRDHPMAEPLPEPKKRLDAQSGSFPFWQOQTRV
BASE COUNT 902 a 1328 c 1182 g 728 t
ORIGIN
Query Match 99.6%; Score 4118.8; DB 74; Length 4140;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 4131; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
Qy 1 taagccccccaccccccgctcagggggtccggttcaggttccttgccttccttcctt 60
Db 8 TAAGCCCCCCCCACCCCGAGCCCGTCCAGGGGCTCCGGTTCAGTTCCTTGCCTCTT 67
Qy 61 cccacgcgcagccttcgc 120
Db 68 CCCACCCGACGCTCTCCGCGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 127
Qy 121 tgggtcgctcagaggaagagtgctcccgggggggaacagtcagtcagtcagtcagtcag 180
Db 128 TGGGTGCGTCAGAGGAGGAGAGTGTC-CCCGGGGGGGAACAGTACAGTACAGTACAG 186
Qy 181 cccgtgcgccttcgc 240
Db 187 CCGCTGCGCGCTCTCCGCGCGCTCCGAGGCTTCCTTCAGCGGAGGCTAAAAGCTCCATC 246
Qy 241 aaacgtacaaagtgcac 300
Db 247 AAACGTACAAAGTCACAAACCAAACTTACCGGACACAGCTTTTCGACAGATCTCGCT 306
Qy 301 cgttcggaagtgcgtgacctgaccgggcccgcgcgcgcgcgcgcgcgcgcgcgcgcgc 360
Db 307 CGCTTCGGAAGTGTGACCATGACCGGCGCGGCTGTGATGCAGAGCTTCAAGGAGTCTCAC 366
Qy 361 tcccatgagtcctgcgtgagtcaccagcagtcgctgctgagccctggagctcaacctggat 420
Db 367 TCCCATGAGTCCCTGCTGAGTCCAGCAGTGTCTGCTGAGGCGCTGGAGCTCAACCTGGAT 426
Qy 421 gaagactccattatcaagccagtcacacagtcacacacacacacacacacacacacacac 480
Db 427 GAAGACTCCATTATCAAGCCAGTACACAGTCCATCTCCGCGGACGAGGATCTGCTTTGAG 486
Qy 481 gtaacacacatgcgtcgggac 540
Db 487 GTAACACATCGTCTGGGACAAATGTTTTCCTGCTCGCTCTGCGAGCGAAGGACACAA 546
Qy 541 tggattgagatctacagagcgctgtgaacacacacacacacacacacacacacacacacac 600
Db 547 TGGATTGAGATCTACAGAGGCGCTGTGAAACCCCAACAGGACACAGCCCGCGGTAGAT 606
Qy 601 aacgtgctgaacatgatcatgaagctcgagcgtgcccccaagacacacacacacacacacac 660
Db 607 AACGTGCTGAACATGATGATCATGAAGCTCGAGGCTGCCCGGCGGCGGCGGCGGCGG 666

Qy 661 tgcagtgatgctgacacatgctctatgacagaccacttccaaagcccccagcc 720
 Db 667 TGCAGATTATGCTTGACGACATGCTATGACAGGACCACTTCAAGCCCGCTAGCC 726
 Qy 721 tcaggaaactatctcttttggggcagacacttcagagtttaacaacctgctgtctcgg 780
 Db 727 TCAGGAGACACTCTCTTTTGGGGCAGACACTTCGAGTTTAAACACCTGCTGTGTCGG 786
 Qy 781 gcgctgcggtgcatctgtacgtgactcgacaaagcgaagcgaagcgaagcgaagc 840
 Db 787 GCGCTGCGGTGATCTGTACCGTGACTCGGACAAAAGCGGAAGAGGACAAAGCGAGGC 846
 Qy 841 taegttagctgactgttccagtgccacctggtggcgccacttcaacagacag 900
 Db 847 TACGTTGGCTGGTGACTGTTCAGTGGCCACCTTGGCTGGCGCCACTTTCACAGAGCAG 906
 Qy 901 tgggtacccctgacctgcaacaggaagtggggctgtgggttatgggtcgggggga 960
 Db 907 TGGTACCCCTGACCTGCAACAGGAAGTGGGGCTCTGGGGTATGGGCTCGGGGGGA 966
 Qy 961 gaggggggtcaggggcggtcaggggcgaagggcgaaggaagagagactcctgtgtcgg 1020
 Db 967 GGAGGGGGTTCAGGGGGGGCTCAGGGGGCAAGGGAAGAGGGCTGTCTGTGTCGG 1026
 Qy 1021 ctgaagggccgttaccagacaatgagtatcctggcccatggagctatatgaaggtttgca 1080
 Db 1027 CTGAAGGCCGCTTACCAGACAATGATATCTTCCCATGGAGCTATATAAGGAGTTTGCA 1086
 Qy 1081 gaatatgtgaccaaacactaccagcatgctgtgctgctgctgagccgcctcaatgtc 1140
 Db 1087 GAATATGTGACCAACCACTTACCGCATGCTGTGCTGGAGCCCGCCCTCATATGC 1146
 Qy 1141 aaggggcaggaggtgctgactgactgactgactgactgactgactgactgactgactgact 1200
 Db 1147 AAGGGCAAGGAGAGGTGCTGTAGTGCATCTGTTTCATCTCTGCAAGGACAGGCAAGGCC 1206
 Qy 1201 aaggactctcttcaacatlggcatgctcagaggttagccgttccatggagcgggaacac 1260
 Db 1207 AAGGACTTCTTTCAGACATGGCATGTCTCAGAGGTAGACCGGTTTCATGGAGCGGGAAC 1266
 Qy 1261 ctcatatccgagagaaacgctcgcactaagcctaaagcctatagaaggtatagaactgatt 1320
 Db 1267 CTCATATTCGCGAGAACACGCTCGCCACTTAAGGCCATAGAAGAGTATATGAGACTGATT 1326
 Qy 1321 ggcagaaatcacctcaagagtgccattggggagttccatccgggctctgtatgaatctgag 1380
 Db 1327 GCCCAGAAATACCTCAAGATGCAATTGGGGAGTTTCATCCGGGCTCTGTATGATCTGAG 1386
 Qy 1381 gagaactgaaatgagacccatcaagtgcacagcgtccagctcgtgagagcaccagcc 1440
 Db 1387 GAGAACGTGAAGTAGACCCCATCAAGTGCACAGCGTCCAGTCTGGCAGACACAGGCC 1446
 Qy 1441 aacctcgagatgctgtgagttggccctgtgcaagggttggtcaactccattgctgttc 1500
 Db 1447 AACCTGGGATGTGCTGAGTTGGCCCTGTGCAAGGTTGTTCACTCCCATTTGGCTTTC 1506
 Qy 1501 ccaggagagctgaagagagttttgcatcatgagcagctgcactatgagagcggggcgg 1560
 Db 1507 CCGAGGAGCTGAAGGAGGTGTTTGGATCATGCGCGCTGCGCTGTGCAGAGCGGGGGCGG 1566
 Qy 1561 gaggacattgctgacaggtgtatcagcgctcgctctctctcgtctcgtctcgtctcgtctc 1620
 Db 1567 GAGGACATGTGTCAGAGGCTGATCAGCGCTCGCTTCTCTCTCTCTCTCTCTCTCTCTCT 1626
 Qy 1621 atcatgtcgccagctctgtttgactgagtagagtagtaccagatgagagacactcacga 1680
 Db 1627 ATCATGTGCGCCAGTCTGTTTGGACATGATGAGAGAGTACCCAGATGAGCAGACCTCAGA 1686
 Qy 1681 accctaccctcatcgcaaggttatccagaacctggccaacttttccaaqttttacatca 1740
 Db 1687 ACCCTCACCTCATCGCAAGGTTATCCAGAACTGGCCAACTTTTCCAAAGTTTACCTCA 1746
 Qy 1741 aaggaggactcctcgggcttcagacgagatttctggagctggaggttctctatgacg 1800

Db 1747 AAGGAGACTTCTCGGCTTCATGAACGAGTTCTGGAGCTTGAATGGGTTCTATGCAG 1806
 Qy 1801 caattcttgatgagatatccaaacctggacacactgacaaacagcagcagtttttgaagggc 1860
 Db 1807 CAATTCTGTATGAGATATCCAAACCTGGACACACTGACCAACAGCAGCAGCTTTTGAAGGC 1866
 Qy 1861 tacatgactgtggcgagctctccacacttccagcctctcgtctgctggaggtgtctgccc 1920
 Db 1867 TACATAGACTTGGCGCGGAGCTTCCACACTTACGCGCTGCTCTGGAGGTGCTGCC 1926
 Qy 1921 caactcaagaagacccctcctgaagctggcgctgcccgcctcctcaacgacatc 1980
 Db 1927 CAGCTCAGCAAGGAACCCCTCTGAAGCTGGCGCCGCTGCCCCGCTCTCAGCGACATC 1986
 Qy 1981 agcacgacctggaggaaaccttaacatccaaaggcagcggcccgagcagcgagcgct 2040
 Db 1987 AGCACAGCCCTGAGGAACCTTAACATCCAAAGGCGAGCGCGCCGAGCGAGCGCGCT 2046
 Qy 2041 cggctcagccctgctgtgctgctggcgctcagccgagatgagggcgtacatgacg 2100
 Db 2047 CGGTCTCAGCCCATGCTGTGCGCGCGCTCAGCGGAANTGCGAGGCTACATGATCGG 2106
 Qy 2101 gacctcaacagctccatgacacttcagcttcattctgctgctgagcctcaacagcctatg 2160
 Db 2107 GACCTCAACAGCTTCCATGACCTTTCAGTCTTTCATGGCTCGAGGCTTCAACAGCTCTATG 2166
 Qy 2161 gacatgctcgtcctccctcccaacaaagagaaaccccgccgcccctcccggtgg 2220
 Db 2167 GACATGCTCGCTCCCTCCCAACCAAGAGAAACCCCGCCGCGCCCTCCCGGTGG 2226
 Qy 2221 ggtaaagacctgttctatgtgagcggcccaacttggcccggtcctcccaagcactatgc 2280
 Db 2227 GGTAAAGACCTGTCTATGTGAGCGCGCCACCACTGGCCCGGTCTCTCCCGAGCATATGC 2286
 Qy 2281 acgagcgtcgggacatcacagagcgggagagagatgctgagtgctcaacaaagtg 2340
 Db 2287 ACGAGCAGCTCGGACATCACAGAGCGGAGAGAGAGTGTGAGTGTCAACAGAGTGTG 2346
 Qy 2341 tcaatgctgacctcagggcagggcctggggcgcccttaacagcagtagtttcc 2400
 Db 2347 TCCATGCTGACCTGAGGGCGACGGGCTGGGGCGGCTTAAACAGCAGTAGTGTGTTC 2406
 Qy 2401 aacctggcagctgttgggacctgttgcactcaagccaggcttcaactgacagcagcctg 2460
 Db 2407 AACCTGGCAGCTGTGGGACCTGTGACCTCAAGCCAGGCTTCACTCAGCAGCAGCTTG 2466
 Qy 2461 ggggtcgccctgacctcccgccgctctcccaagagagtgctcttccatccagca 2520
 Db 2467 GGGTGGCGCTGCACTTCCCGCGGCTCTCCCAAGGGAGTGGCTCTTCATCAGAGCA 2526
 Qy 2521 gccggcatgctcagccagatgggtgtcactcaggtggtgtgtcccgcccgcaactg 2580
 Db 2527 GCCGGCATGCGCTTACGCGAGATGGGTGTGCTACGAGGATGGTGTGCCCGCCAGCAACTG 2586
 Qy 2581 cgcacctcttctccagaaacctcttccatagctgctgctgagtgagcagggggcc 2640
 Db 2587 CGCATCTCTTCTTCTTCCAGAACCTCTCTTCATATGGCTGCGGATGGAGGCGCC 2646
 Qy 2641 ccagcagccatggaggagcagtggttccaccttctccatccatccacacac 2700
 Db 2647 CCAGCAGGCTTGGAGGAGCAGTGGCCATGTGTCACCTTCTCCCATCAGCAGCAGCAC 2706
 Qy 2701 caccatcacctacagggggagaaaccccgagggacacttttggccggttccatggc 2760
 Db 2707 CACCATCACCATCAGGAGGGGAGAACCCCGCAGGGGACACTTTTGGCCCGTTCATGGC 2766
 Qy 2761 tatagcagcagcagcactctctacaggggttccctaaagccctcgccctccatcctt 2820
 Db 2767 TATAGCAAGAGCAGGAGACCTCTCTACAGGGGTCCCTAAAGCCCTTGGCGCTTCCATCTT 2826
 Qy 2821 cacagccacagctacagtagtggagcttggacctctgttactgattttaccctgagcag 2880

QY	3961	gaagctcaacaaaggagacagcgaggattgtcttccttaacgcctctcttggggcaaccatctg	4020
Db	3967	GGAGCTTCACCAAGGAGACAGCGGGATGTCTCCCTTAACGCCCTCTTGGGGCACACCATCTG	4036
QY	4021	tcaacccactgctcatcttcagagggagagtgggaccctcagctgcctctcacccca	4080
Db	4027	TCAACCCCACTGCTCCATTCACGAGGAGAGTGGGACCTCAGCTGCCTCTCACCCCA	4086
QY	4081	ggaacacactatcccccacacagacccttcactctgggtgtatcccccactct	4134
Db	4087	GGACACCACTTACCCACACAGACCCCTTCACCTCTGGGTGCTATCCCACTCT	4140

2

RESULT	AF050183	4265 bp	mRNA	ROD	15-SEP-2000
LOCUS	Rattus norvegicus GTPase activating protein Syngap-c mRNA, complete cds.				
DEFINITION	AF050183.2 GI:10140854				
ACCESSION	AF050183				
VERSION	1				
KEYWORDS	Norway rat.				
SOURCE	Rattus norvegicus				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.				
REFERENCE	1 (bases 1 to 4265)				
AUTHORS	Kim, J. H., Liao, D., Lau, L. F. and Haganir, R. L.				
TITLE	Syngap, a synaptic RasGAP that associates with the PSD-95/SAP90 protein family				
JOURNAL	Neuron 20 (4), 683-691 (1998)				
MEDLINE	98240917				
REFERENCE	2 (bases 1 to 4265)				
AUTHORS	Kim, J. H. and Haganir, R. L.				
TITLE	Direct Submission				
JOURNAL	Submitted (20-FEB-1998) Neuroscience, Johns Hopkins Medical Institute, 725 N. Wolfe Street, PCTB 900, Baltimore, MD 21205, USA				
COMMENT	On Sep 15, 2000 this sequence version replaced gi:3044054.				
FEATURES	Location/Qualifiers				
source	1. .4265 /organism="Rattus norvegicus" /db_xref="taxon:10116" /tissue_type="hippocampus" /dev_stage="adult"				

```

477..3792
/note="neuronal synaptotagmin RasGAP: associates with the
PSD-95/SAP90 protein family; alternatively spliced
variants, SynGAP-a and SynGAP-b"
/codon_start=1
/product="GTPase activating protein SynGAP-c"
/protein_id="AAC40082.2"
/db_xref="GI:10140855"
/translation="MQSFKESSHSHESLLSPSSAAFALELNDEDSILIKPVHSSILGQE
FCEVTVISSGTTKFCACRSAAERDKRIENLQRAVPKPNRRDVNVLKMLILAEBSLP
PKKRYICELCIDMLLARTISIKPRSSAQDTVWGEHFEFTFNRLPAVRALRIKLRDSDP
KRGKGGAVGLVTVPVATLGRHFEHQWYPTLPTGGSGGGSGGGSGGGSGGGSGG
KRGKGGCAVAKLARYGTQSLIPMELEYKFAEYVYTNHMLCAVLEPAAVNLKGEVAS
ALVHIQSTKALDEKLSDMMAEVEQRFMERHEHLIRNTLAKTALIEYMLIKQRYLKL
DAIFGTGKALYESEENCVDVPTKCFATSSALHQAQNMCCCELALCKVNVSHCVQRL
KEFVAFNRLEACBREDGIDADRLIISASLFLFCFPAIMSPSLFLGIMDEYDEQSTSTL
TLIAKVIOLANFSKTSKEDFLPQMFLEFGSMQOFLYEINLDTLTNDSFEG
YIDLGRPLTSLHAALMEVLPOLSKALKILGLPRLILSDTSLARNPNIQORPSROSE
RARSQWPLTSGSAWQGYMRDLNLSLIDLOSMARGINSSMDMARLPSTPEKRPFP
PRGGGLDGVPSRPAKSSPAYCTSSDITPEQPKMLSYNKSVMILDLOGDGGGGL
NSSVSNLNAVGLLHSSQASLTALGLRPAAGRLSQGSGSSLTAAHMLRSONGVTT
DGVFAQQLRPLSFGNPLFHMAADGGPPAGHGSGSGHPPSSHHHHHHHHHHRGEP
PGDTFFHFGYSKSDSLGTGVPAPASILHSHSYSDFFPGSGDTFTRLQSLQDMLQ
HMLSPQTITGPORAPSGPGGGGGGGGGPPPLQKSGQSLTYSAAQKPRPS
SNGLASPEYSGYPARPOQSLSKEGSTGGSGGGGGGGLPSTLTKHQSPTPLS
PTMPSERTVAVYNNHPIASDIESAHIERBEYKIKYISKMSDESLRDYRFEETIH
SLAKERLHMSNRKLEELYSQEQPTSKILMOYQARLEQSEKRLQQQVQKDSQSL
IIGRLMLVEBELRDHPMAEYLPPEPKRLLDAQVROLPLPLGNVPRVLPAPWNLG

```

APPAPPPPRLQITENGFEFRNTADH"
BASE COUNT 936 a 1358 c 1203 g 768 t
ORIGIN

Query Match 94.8%; Score 3919; DB 74; Length 4265;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 3942; Conservative 0; Mismatches 15; Indels 4; Gaps 1;

BASE COUNT	936 a	1358 c	1203 g	768 t	APPAPPPPPRLQITENGEPFRNADH"
ORIGIN					
Query Match	94.8%	Score 3919;	DB 74;	Length 4265;	
Best Local Similarity	99.3%	Pred. No. 0;			
Matches 3942;	Conservative	0;	Mismatches 15;	Indels 4;	Gaps 1;
QY	178	gcccgcgtgcccctccggccctgcagcaggtctctctgagccggagggtctaaagctcc	237		
DB	118	GTCAAGGGCAAGGAGGAGGTGCTAGTGCATTCCTGCAATTCCTGCAAGAGCAAGGCAAG	1137		
QY	238	atcaaacgtacaaagtacaaacccaaacttgaccggaccagagctttcgacagatccgtg	297		
DB	178	ATCAAAAGGTACAAATCAACACCAACTTGACCGGACCAAGAGCTTTCGACAGATCTCTG	237		
QY	298	ctctgccttcgaagtgctgaacatgaccggggccggctgctgacagagcttcaaggagctc	357		
DB	238	CCTCGCTTCGGAAGTGCCTGACCATGACGGGCGCGGTGATGACAGAGCTTCAAGAGTCT	297		
QY	358	cactcccatgagtcctcgtcgtgagtcaccagagtgctgtaggcccctgagctcaaacctg	417		
DB	298	CACCTCCCATGAGTCCCTGCTGAGTCCCAAGCAGTGTCTGCTGAGGCCCTGGAGCTCAACCTG	357		
QY	418	gatgaagactccattacaagccaggtacacagctccatcctgagccagaggttctgcttt	477		
DB	358	GATGAAGACTCCCTTATCAGCCAGTACACAGCTCCATCTCTGGCCAGAGGATTCGTCTTT	417		
QY	478	gaggttaacacatcgtctgggacaaatgttttgcctgctcgtgtcagccgaaaggagac	537		
DB	418	GAGGTAAACAATCTGCTGGGACAAAATGTTTTTGGCTGTGCGGTCTGCAGCCGAAAGGAG	477		
QY	538	aaatgattgaaatctacagaggtctgaaaccccaagacacacagcgcgcggtga	597		
DB	478	AAATGATGAGAACTACAGAGGGCTGTGAACCCCAAGAACACACAGCCGCCGGGTA	537		
QY	598	gataacgtgctgaaactatgataagaagctgagagctgcccccaagaagcgatat	657		
DB	538	GATAACGTCTGAAACTATGATATAGAAGCTGCAGAGGTGCCCCCACAAGAGCGGATAT	597		
QY	658	tactgcgagttatgcttgagcagacatgctctatgcagcagccacttccaaagccgcctca	717		
DB	598	TACTGCGAGTTATGCTTGAGCAGACATGCTCTATGCAGCGACACTTCCAAGCCCGCTCA	657		
QY	718	gctcaggagacactgtcttttggggcgagcaacttcagtttaacaaactgcctgctgctc	777		
DB	658	GCTCAGGAGACACTGTCTTTTGGGGCGAGACATTCAGATTTAACAACTTGCCTGCTGTC	717		
QY	778	cgggcgtcggtcgtcatctgtaccgtgactcgagcaaaaagcgggaaggacaaaggca	837		
DB	718	CGGGCGCTCGGCTGCATCTGTACCGTGTGACTCGACAAAAAAGCGGAAGGACAAAGCA	777		
QY	838	ggctaagttggcctgagtgactgttccagtgccacactggctgggcccacttcaagag	897		
DB	778	GGCTACGTTGGGCTGTGACTGTTCAGATGGCCACCCTGGCTGGGCGCACTTTCACAGAG	837		
QY	898	cagtgatccccgtgacctctgcaacagaaagtgggggctcgtggggatgggctcgggg	957		
DB	838	CAGTGTACCCCGTGACCTTGCAACAGAGAAAGTGGGGGCTCTGGGGGTATGGGCTCGGG	897		
QY	958	ggaagagggggtcaggggcggtcgaaggggcaaaaggaaaggaggtctcctgctgtg	1017		
DB	898	GGAGGAGGGGGGTCAAGGGGGCGGCTCAGGGGGCAAGGAAAGAGAGGCTCTCTCTGTGTG	957		
QY	1018	cggtgaagggccgttaccagaaatgagtatcctgcgcctgagctataaaggagttt	1077		
DB	958	CGGCTGAAGGCCGCTTACCAGACAATGATATCTGCCCATGGAGCTATATAAGGAGTTT	1017		
QY	1078	gcagaatatgacaaacactacagatgctgtgcgtgagccgcctccaat	1137		
DB	1018	GCAGAATAATGTGACCAACCACTTACCGCATGCTGTGTGCGGTGTGTGGAGCCGCCCTCAAT	1077		

Db 2158 GGGGGTAAAGACCTGTTCTATGTGAGCGCCACACACTGGCCCGCTCTCCCCAGCATAC 2217
 Qy 2278 tgcacgacgactcgacatcacagagccgagcagaaga tgcctgagtgatcaacaagagt 2337
 Db 2218 TGCACGAGCAGCTCGGACATCATCAGAGCGGAGCAGAGAGATGCTGAGTGTCAACAAGAGT 2277
 Qy 2338 gtgtccatgctgacatcgagggcagcggcctgagggcgccttaacagcagtagtggtt 2397
 Db 2278 GTGTCCATGCTGGACCTGCAGGCGCAGCGGCTGGGGCGCGCTTAACAGCAGTAGTGT 2337
 Qy 2398 tcaaacctggcagctgttgggacactgttgcaactcaaacgaagcttcaactgaagaagcc 2457
 Db 2338 TCAAACTGGCAGCTGTGGGACCTGTGCACTCAAGCCAGGCTTCACTGACAGCAGCC 2397
 Qy 2458 ttgggggtggcgcctgacactccggggcgctctcccaaggagtggtgctcttccatcaaca 2517
 Db 2398 TTGGGGTGGCGGCTGCACCTGCGGGCGGCTCTCCCAAGGGAGTGGCTCTCCATCACA 2457
 Qy 2518 qcagccggcagctgacagcagatgggtgtcaactcagatggttcccgccagcaaa 2577
 Db 2458 GCAGCGGCATGCGCTGAGCCAGATGGGTGTCACTACGGATGGTGTCCCGCCAGCAA 2517
 Qy 2578 ctgagcaatccctttctctccagaaacctctcttccatatagtgtgcgatgacaaaggg 2637
 Db 2518 CTGGCATCTCTTTCTTCCAGAACCTCTCTTCCATATGGCTGCGGATGGAGACAGG 2577
 Qy 2638 ccccaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2697
 Db 2578 CCCCCAGCGCCATGAGGAGCAGTGGCCATGTCACCTCTCTCCATCACCACAC 2637
 Qy 2698 caccacatcaccatcaccagggagagagagagagagagagagagagagagagagagagag 2757
 Db 2638 CACCACCATCACCATCAGCGAGGGAGAGAACCCCGAGGGAGACATTTGGCCCGTTCAT 2697
 Qy 2758 ggcctatagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2817
 Db 2698 GGCCTATAGCAGAGCAGGAGCAGCTCTCTACAGGGGTCCCTTAAGCCCGCTGGCGCTCCATC 2757
 Qy 2818 ctccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2877
 Db 2758 CTTACAGCCACAGCTACAGTGTGATGTTGGACCTCTGCTACTGATTTTACCGCTCG 2817
 Qy 2878 cagctctcaactcagagaaactcaagacatgctctcccgccccagagacacacatcaggt 2937
 Db 2818 CAGCTCTCACTTACAGAACACCTTACAGCAGATGCTCTCCCCCGCCAGATCACCATCGGT 2877
 Qy 2938 ccccaagggcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2997
 Db 2878 CCCCAGAGCCAGCTCCCTCAGGGCCAGGAGGGGCGAGTGGTGGGGCGAGTGGTGGGGG 2937
 Qy 3057 ggtggggcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3117
 Db 2938 GGTGGGGCCAGCCACTCCCTTGCAGAGGGGCAATCTCAGCAGTTGACAGTGAAGTGT 2997
 Qy 3117 gccagaaaaacccggcgttcagcagggaaacctattcagctcccgccggaacaaagttagt 3177
 Db 2998 GCCCAGAAACCCCGGGCGTCCAGCGGGAACTTATTCAGTCCCCCGAACCAAGTTATGGT 3057
 Qy 3178 cctgcccgtccacggaacacagagcctcagcaagaggcagcagcagcagcagcagcagcagc 3177
 Db 3058 CTTCCCGCTCACGGCAACAGAGCCCTCAGCAAAAGAGGGCAGCATTTGGGGCGAGCGGGG 3117
 Qy 3178 agcgggtggcagggggtggggggtcagcagcagcagcagcagcagcagcagcagcagcagc 3237
 Db 3118 AGCGGTGGCGAGGGGTGGGGGTCTCAGCCCTCCATCACCAGCAGCATTTCCAGACT 3177
 Qy 3238 ccataccagctgaacccccagatgcggcctcgagcagcagcagcagcagcagcagcagcagc 3297
 Db 3178 CCATCCAGCTGAACCCAGATGCGCGCTCGGAGCGGAGTGTAGCTGGGTGTCCAAAT 3237
 Qy 3298 atgcctcacctgtccgtgacatcagagtgacacatgagcagcagcagcagcagcagcagc 3357

Db 3238 ATGCTCACCTGTGCTGACATCGAGAGTGACACACTTGGCGGAGAGTACACAGCTG 3297
 Qy 3358 aaggagttactcgaagtcactgtgacagagagcagcagcagcagcagcagcagcagcagcag 3417
 Db 3298 AAGAGTACTCTGAAGTCCATGGACGAGAGCCAGCTGGACAGGGTGAAGAGTACTAGAGG 3357
 Qy 3418 gagatccactcactgaaggaaggtcacacatgtccaaccggaagctggaagagtagcag 3477
 Db 3358 GAGATCCACTCACTGAAGGAAGGCTACACATGTCCAACCGGAAGCTGAAGAGTACGAG 3417
 Qy 3478 cggagctgctgtccagagaagagcagcagcagcagcagcagcagcagcagcagcagcagc 3537
 Db 3418 CGGAGGTGCTGTCCAGGAGAGCAGCAGCAAGATCTCTGATGACATGACAGCCCGC 3477
 Qy 3538 ctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 3597
 Db 3478 CTGGAGCAGAGCAGAGAGGCTTTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 3537
 Qy 3598 agatcatctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 3657
 Db 3538 AGCATCATTTGCGAGGCTGATGCTGTGTGAGAGGAGCTGCGCCGGGACACCCCGCATG 3597
 Qy 3658 gctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3713
 Db 3598 GCTGAGCGCTGCTGAACCAAGAGAGGCTGCTGAGCAGCTCAGGTGGAGAGCAGCTT 3657
 Qy 3714 cccccctgggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 3773
 Db 3658 CCCCCCTGGGTGCTCAACAAACCCGCTGTGAGCTGGCGCCACCTTGGAAAGCGCTGGCC 3717
 Qy 3774 cccccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3833
 Db 3718 CCCCCAGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCC 3777
 Qy 3834 accgacacacactagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 3893
 Db 3778 ACCGACAGCAGCTAGCGCCAGCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 3837
 Qy 3894 cccacccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3953
 Db 3838 CCCCAGCGGTGACAGAGCTCTTCCAGTGCACCCAGCCAGCCAGCCAGCCAGCCAGCCAG 3897
 Qy 3954 cagagctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 4013
 Db 3898 CAGAGCTGGAGCGTACCAAGGAGACAGCGGATGCTCTTCCCTTAAGCGCTCTTGGGAC 3957
 Qy 4014 ccactgtcaacccactgctccattccagggagggagggagggagggagggagggagggaggg 4073
 Db 3958 CCATCTGTCAACCCCTGCTCCATTTCCAGAGGAGGAGTGGGACCCCTCAGCTGCCCTCT 4017
 Qy 4074 caccacggagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 4133
 Db 4018 CACCCAGGACACACCTACCCACACACAGCCCTTCACTCTGGGGTGTCTATCCCATCC 4077
 Qy 4134 t 4134
 Db 4078 t 4078

RESULT AF058789 4539 bp mRNA ROD 14-SEP-2000
 LOCUS Rattus norvegicus SynGAP-a mRNA, complete cds.
 DEFINITION AF058789
 ACCESSION AF058789.2 GI:10122137
 VERSION AF058789.2
 KEYWORDS Norway rat.
 SOURCE Rattus norvegicus
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (bases 1 to 4539)
 AUTHORS Kim, J.H., Liao, D., Lau, L.F. and Haganir, R.L.

QY 1491 ttgcgtgttcccgaggagctgaagggtgtttgcatcatgagcggtcgcgtgtgcaga 1550
 Db 1629 TTGCGGTGTCCCGAGGAGCTGAAGAGGTGTGTTCATATGCGGCTGCGTGTGCAGA 1688
 QY 1551 gggggccggaggacattgctgacagctgatcagcctcgtcttctcgtctct 1610
 Db 1689 GCGGGCCGGGAGGACATTTGCTGACAGGCTGATCAGCGCTCGCTCTTCTCGGCTTCT 1748
 QY 1611 ctgcccgcacatgctgcgccagtctgttgactgatcagagataccagatgagca 1670
 Db 1749 CTGCCCCGACATATGTCGCCAGCTGTGTGACATGATCAGAGATACCCAGATGAGCA 1808
 QY 1671 gacctcagaacacctcaccctcatcgccaagggttatccagaacctggccaacttttccaa 1730
 Db 1809 GACCTCAGAAACCTCAACCTCATCGCAAGGTATTCAGAACTGGCCAACTTTTCCAA 1868
 QY 1731 gtttacctcaaggagagacttctggcttcatacgaagtttctgagctgagatggg 1790
 Db 1869 GTTTCACCTCAAGAGGAGACTTCTTGGGCTTCATGAAGAGTTTCTGGAGCTGAGTGGG 1928
 QY 1791 ttctatgcagcaattcttctgatgatataccaaacctggacacactgaccaacagcag 1850
 Db 1929 TTCTATGAGCAATTTCTGTATGATATCAACCTTGGACACACTGACCAACAGCAGCAG 1988
 QY 1851 ttttgaggtctacatgacttggccgagctctccacacttcagccctcgtctgga 1910
 Db 1989 TTTTGAGGGCTACATAGACTTGGGCCCGGAGCTCTCCACACTTACGCCCTGCTTGGGA 2048
 QY 1911 ggtctgccccagctcagcaaggagccctctcctgaagctgggcccgtgccccgtct 1970
 Db 2049 GGTGCTGCCCGCAGCTCAAGCAAGGAGCCCTCTCTGAAGCTGGGCGCTGCTCCGCTCT 2108
 QY 1971 cagcgacatcagcagccctgaggaaccttaacatccaaaggcagccgagccgacag 2030
 Db 2109 CAGCGACATCAGCAGCCCTGAGGAGACCTTACATCCAAAGGACCGGAGCGCCAGAG 2168
 QY 2031 cagagcgctcgttctcagccatgtgtgctgctgctgctgctgctgctgctgct 2090
 Db 2169 CGAGCGCGCTCGGCTCAAGCCATGTGTGCTGCGCGGCGCTCAGCGAGATCAGGGGTA 2228
 QY 2091 catgatcgggagctcgaacagctccatcgacaccttcagctctcctcgtcagagctcaa 2150
 Db 2229 CATGATCGGGACCTCAACAGCTCCATCGACCTTTCAGTCTCTTTCAGTCTGAGGCTCAA 2288
 QY 2151 cagctctatgacatgctgctcctcctcctcctcctcctcctcctcctcctcctcct 2210
 Db 2289 CAGCTCTATGAGACATGCTCGCTTCCCTTCCCTTCCCAACCAAGAGAAACCCCGCGGCCCC 2348
 QY 2211 tcccgggtggggtaaagacctgttctatgtgagccggccacctggtgctcctccc 2270
 Db 2349 TCCCGGTGGGGTAAGACCTGTTCATGTGAGCGCGGCCACCACTGGCCCGGTCTCTCCOC 2408
 QY 2271 agcatactgcagcagctcggacatcacagagccgagcagagatcagatgtcaa 2330
 Db 2409 ACATACTGACAGAGCTCGGACATCACAGAGCCGAGCAGAAATGCTGAGTGTCAA 2468
 QY 2331 caagagtgtccatgctggaacctgcaggcgagcgctggggcgcttaaacagcag 2390
 Db 2469 CAAGAGTGTTCATGTGGACCTGCAAGCGAGCGAGCGCTGGGGCGCGCTTAACAGCAG 2528
 QY 2391 tagttttccaaacctgcagctgttgggagctgttgcactcaagcagagcttactgac 2450
 Db 2529 TAGTGTTCACACCTGGCGAGCTGTGGGAGCTGTGACATCAAGCGAGGCTTTCATGAC 2588
 QY 2451 agcagccttggggtgcccgtgcacctgcggcgctctcccaaggagtggtcttc 2510
 Db 2589 AGCAGCCTTGGGTTGGCGCTGACCTGCGCGCGCTCTCCCAAGGAGTGGGTCTTC 2648
 QY 2511 catcacagcagcagctgctcagcagcagatgggtgtcactcagatgggtgtccgcg 2570
 Db 2649 CATCACAGCAGCGGATGCGGCTTCACCCAGATGGGTGTCTACTACGGATGGGTGTCGCCG 2708
 QY 2571 ccagcaactgcagctcctcttcttccagaaacctctcttccatgtggtgcagatgg 2630

Db 2709 CCAGCAACTGCGCATCCCTCTTCTTCAGAACCTCTCTTCATATAGGTGCGCATGG 2768
 QY 2631 accagggcccccagcagggccatggaggagcagtgagccatggtccaccttctctccca 2690
 Db 2769 ACCAGGGCCCCAGCAGCGCATGGAGGAGCAGTGGCCATGCTCCACCTTCTCCCATCA 2828
 QY 2691 ccacacccaccacatccacatccagcaggggagaaaccccccagggacactttggccc 2750
 Db 2829 CCACACACACACATCACCATCACCGAGGGGAGAACCCCGAGGGACATTTTGGCCCC 2888
 QY 2751 gttccatggtctatagcaagcagcagcctctctacaggggtccctaaagccccctgcgc 2810
 Db 2889 GTTCCATGCTATAGCAAGAGCGAGACCTCTCTACAGGGTTCCTTAAGCCCCCTGCGGC 2948
 QY 2811 ctccatctccacagccacagctacagtgatgagtttgagccctctgctgactgatttac 2870
 Db 2949 CTCCATCTTCACAGCCACAGCTACAGTATGAGTTTGGACCCCTCTGCTACTGATTTTAC 3008
 QY 2871 cctgctgcgcgtctctcacttcaggacacactcagcagcactgctctcccccgcagatcac 2930
 Db 3009 CCCTCGCGAGCTCTCACTTCAGGACAACTACAGACACATGCTCTCCCGCCCGACATCAC 3068
 QY 2931 catcgtccccagagccagctccctcagggccaggggagggcagtggtggggcagtg 2990
 Db 3069 CATCGGTCCCCAGAGGCCASCTCCCTCAGGGCCAGAGGGGGCAGTGGTGGGGCAGTGG 3128
 QY 2991 tggggcggtggggggcagccacctctctgagagggggcaaatctcagcagtggtgacagt 3050
 Db 3129 TGGGGCGGTGGGGGCGAGCCACCTCCCTTTCAGAGGGGCAAACTCTCAGCATGTTGACAGT 3188
 QY 3051 gactgtgcccaagaaaccccgccgtccagcaggggaacctattgagtcctcccgagcaacaag 3110
 Db 3189 GAGTGTGCCCGAGAAACCCCGCGCTCCAGCGGGAACCTATTGCACTCCCGGGAACCAAG 3248
 QY 3111 ttatgctcgtcccgctccacggcacaagcagcctcagcaaaaggggcagcattggggcag 3170
 Db 3249 TTATGTCTTGGCCGCTCCACGGCAACAGAGGCTCAGCAAAAGAGGGCAGCATTTGGGGGCGAG 3308
 QY 3171 cggggcgagcgggtggcgagggggtgggggggtcgaagccctccatccacaaagcagcttc 3230
 Db 3309 CGGGGCGAGCGGTGGCGGAGGGGTGGGGGGCTCAAGCCCTCCATCACCAAGCAGCATTC 3368
 QY 3231 ccagactccatccagctgaaccccaacacacacacacacacacacacacacacacacacac 3290
 Db 3369 CCAGACTCATCCACCTCAACCCACGATCGCGCTCGAGGCGGACTGTAGCTCGGT 3428
 QY 3291 gtcccaatagcctcactgctcgcgtgacatcgagagtgacacacacacacacacacacacac 3350
 Db 3429 GTCCAAATATGCTCACTGCTCGCTGACATCGAGAGTGCACACATTTGAGCGGGAAGATA 3488
 QY 3351 caagctgaaggagctactcgaagctccatggacagagcagcactgacacaggtggaaggata 3410
 Db 3489 CAAGCTGAAGGAGTACTCGAAGTCCATGGACGAGAGCCGACTGGACAGGGTGAAGAGATA 3548
 QY 3411 cgaggaggagatccactcactgaagaaaggctacacatgtcccaacccggaagctggaaga 3470
 Db 3549 CGAGGAGGAGATCCACTCACTGAAGAAAGGCTACACATGTCCAACCGGAAGCTGGAAGA 3608
 QY 3471 gtacagcagagctcgtctctccagaaagacagaccagcagatcctgagtcagatccca 3530
 Db 3609 GTACGAGCGAGGCTGCTCTCTCCAGAGAGCAGACCAAGATTCCTGATGCGATACCA 3668
 QY 3531 agccccctggagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3590
 Db 3669 AGCCCCCTGGAGCAGAGCGAGAGCGCTTTCAGGCGAGCAGAGGTGGAGAGGACTCCCA 3728
 QY 3591 gatcaagacatcattgagcagctgactgagagagagcagcagcagcagcagcagcagcagcagc 3650
 Db 3729 GATCAAGACATATTGGCAGGCTGATGCTGTGTGAGGAGGAGTGTGCGCGGAGCACCC 3788
 QY 3651 cgccatggtgagcgcgtcctgaaaccccaagaggcgtgctcgacgctca----- 3701

Db	3789	CGCCATGGCTGAGCGCTGCCTGAACCAAGAGAGGCTGCTGACGCTCAGCTCCTCAT	3848
QY	3702	-----	3701
Db	3849	CAGGTAAATTCTCCTGGTTCCCGCTTTGACACACGGGGGAGGACACTGGGGGAGGTGACTCC	3908
QY	3702	-----gagagcgagcttcccccttgggtccaaacaaacccgcgtgacgcctg	3749
Db	3909	GGACCACATGCAGGAGAGGCAGCTTCCCTCCCTTGGGTCCCAAAACCCGCGTGTGACGCTG	3968
QY	3750	gccccaccttgaaacgcctggtgccccccagccccccccccccccccccccccccccgcgtgcagatc	3809
Db	3969	GCCTCCACTTGGAAAGGCCTGCGCCCGCCAGCCCCACCCCCCCCACCGCGGTGCGAGATC	4028
QY	3810	acagaaacgcgcgagttccggaacacccgcagaccactagccacccagcatcacagacct	3869
Db	4029	ACAGAAACCGGGGAGTTCCGGAAACACCGAGACCACTAGCCCCACCCAGCATCACAGACCT	4088
QY	3870	ccttcctctgtgcacctataccccggccccaccagctcacagacctcttcccagtgacc	3929
Db	4089	CCTTCCCTGTGCACCTTACCCGGCCCCACCCAGCGTCACAGACCTCCTTCCAGTGCACC	4148
QY	3930	cgaccttgaaacatcaccaacacacagagactggaogtcaccaaggagacagcggaattgctc	3989
Db	4149	CGACCCCTGGAACATCACCAACACCAGGACTGGAGCTCACCAGGGACAGCGGATTTGTC	4208
QY	3990	tcctttaagcctctcttgggggaccccatctgcaacccccactgctccattccagagagga	4049
Db	4209	TCCTTTAAACGCCCTTCCTTGGGGGACGG-TCCTGTCAACCCCACTGTCTCATTTCCAGAGGGA	4267
QY	4050	gagtgggaacctcagctgcccctctcacccccaggacacacacctaccaccaacagacccctt	4109
Db	4268	GAGTGGGACCTCAGCTGCCCCCTCACCCCAAGACACCACTTACCCCAACAGACCCCTT	4327
QY	4110	cactctggggtgctatcccatact	4134
Db	4328	CACCTCTGGGTGTATCCCCATCCT	4352

Log to No. 4

RESULTS 4 AF048976 4063 bp mRNA 27-MAR-1998
LOCUS Rattus norvegicus synaptic ras GTPase-activating protein p135
DEFINITION SRYGAP mRNA, complete cds.
ACCESSION AF048976
VERSION AF048976.1 GI:2935447
KEYWORDS Norway rat.
SOURCE Rattus norvegicus
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus

1 (bases 1 to 4063)
Chen, H.-J., and Kennedy, M.B.
Identification and cloning of a novel 130 kd protein containing a
ras GTPase-activating domain from the rat forebrain postsynaptic
density
(in) SOC. NEUROSCI. ABSTR.: 1456; (1997)

JOURNAL
REFERENCE 2 (bases 1 to 4063)
AUTHORS Chen, H.-J., Rojas-Soto, M. and Kennedy, M.B.
TITLE A synaptic Ras GTPase-activating protein (p135 Syngap) inhibited by
Cam kinase II
Unpublished

JOURNAL
REFERENCE 3 (bases 1 to 4063)
AUTHORS Chen, H.-J. and Kennedy, M.B.
TITLE Direct Submission
JOURNAL Submitted (17-FEB-1998) Division of Biology, California Institute
of Technology, 1200 E. California Blvd., MC 216-76, Pasadena, CA
91125, USA

FEATURES Location/Qualifiers
source location .4063
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"

RESULT 5

RESULT	5	ABO16962	4801 bp	mRNA	ROD	01-SEP-2000
LOCUS		ABO16962	4801 bp	mRNA	ROD	01-SEP-2000

DEFINITION	ACCESSION	VERSION	RAIUS NORVEGICUS IMAGE
	AB016962	AB016962.1	GI:4417206

KEYWORDS
SYNGAP-1
PATTUS NORVEGICUS (sub species: Sprague Dawley) cDNA to mRNA.
SOURCE

ORGANISM	Vertebrata:	Crustacea:	Chordata:	Metazoa:	Protozoa:
<i>Rattus norvegicus</i>					
<i>Nauplius</i>					
<i>Paramecium</i>					
<i>Amoeba</i>					
<i>Dictyostelium</i>					
<i>Planaria</i>					
<i>Hydra</i>					
<i>Paramecium</i>					
<i>Amoeba</i>					
<i>Dictyostelium</i>					
<i>Planaria</i>					
<i>Hydra</i>					
<i>Paramecium</i>					
<i>Amoeba</i>					
<i>Dictyostelium</i>					
<i>Planaria</i>					
<i>Hydra</i>					
<i>Paramecium</i>					
<i>Amoeba</i>					
<i>Dictyostelium</i>					
<i>Planaria</i>					
<i>Hydra</i>					
<i>Paramecium</i>					
<i>Amoeba</i>					
<i>Dictyostelium</i>					
<i>Planaria</i>					
<i>Hydra</i>					
<i>Paramecium</i>					
<i>Amoeba</i>					
<i>Dictyostelium</i>					
<i>Planaria</i>					
<i>Hydra</i>					
<i>Paramecium</i>					
<i>Amoeba</i>					
<i>Dictyostelium</i>					
<i>Planaria</i>					
<i>Hydra</i>					
<i>Paramecium</i>					
<i>Amoeba</i>					
<i>Dictyostelium</i>					
<i>Planaria</i>					
<i>Hydra</i>					
<i>Paramecium</i>					
<i>Amoeba</i>					
<i>Dictyostelium</i>					
<i>Planaria</i>					
<i>Hydra</i>					
<i>Paramecium</i>					
<i>Amoeba</i>					
<i>Dictyostelium</i>					
<i>Planaria</i>					
<i>Hydra</i>					
<i>Paramecium</i>					
<i>Amoeba</i>					
<i>Dictyostelium</i>					
<i>Planaria</i>					
<i>Hydra</i>					
<i>Paramecium</i>					
<i>Amoeba</i>					
<i>Dictyostelium</i>					
<i>Planaria</i>					
<i>Hydra</i>					
<i>Paramecium</i>					
<i>Amoeba</i>					
<i>Dictyostelium</i>					
<i>Planaria</i>					
<i>Hydra</i>					
<i>Paramecium</i>					
<i>Amoeba</i>					
<i>Dictyostelium</i>					
<i>Planaria</i>					
<i>Hydra</i>					
<i>Paramecium</i>					
<i>Amoeba</i>					
<i>Dictyostelium</i>					
<i>Planaria</i>					
<i>Hydra</i>					
<i>Paramecium</i>					
<i>Amoeba</i>					
<i>Dictyostelium</i>					
<i>Planaria</i>					
<i>Hydra</i>					
<i>Paramecium</i>					
<i>Amoeba</i>					
<i>Dictyostelium</i>					
<i>Planaria</i>					
<i>Hydra</i>					
<i>Paramecium</i>					
<i>Amoeba</i>					
<i>Dictyostelium</i>					

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

REFERENCE 1 (bases 1 to 4801)

AUTHORS	Suzuki, T.
TITLE	SynGap-d

JOURNAL
REFERENCE
Published Only in Data
2 (bases 1 to 4801)

AUTHORS Suzuki, T.
Direct Submission

JOURNAL TITLE
Direct submission
Submitted (14-AUG-1998) to the DDBJ/EMBL/GenBank databases. Tatsuo
submitted to the school of medicine, Department of
Chiba University, Chiba City, Japan.

SUZUKI, Shinshu University School of Medicine, Department of Neuroplasticity; 3-1-1 Asahi, Matsumoto, Nagano 390-8621, Japan

(E-mail: suzukit@sch.md.shinshu-u.ac.jp, Tel: +81-263-37-2800)
Fax: +81-263-37-2725)

On Mar 16, 1999 this sequence
Sequence updated (09-Feb-1999).

FEATURES

```
source
1. 14001
/organism="Rattus norvegicus"
```

```

/sub_species="Sprague Dawley"
/db_xref="taxon:10116"

```

gene
1124.4624
/gene="synG

```
CDS
1124, .4624
/gene="svnGAP-d"
```

```

/ gene="synaptic ras GTPase activating protein"
/ note="synaptic ras GTPase activating protein"
/ codon_start=1

```

```

/conn_start=
/product="synGAP-d"

```

```

/protein_id="BAA749"
/db_xref="GI:441720"

```

```
/translation="MEYFPPAAPFPSPQGFSLSRKRLKSSIAKRTASQPKLDRYSSFRQLLF  
RFRSADHDRARLMOQFKESHESHESLLSPSSAAEALELNLEDSIIKPVHSSILGQEFC
```

FEVTTSSGTKCFACRSAAERDKWIENLQRAVKPNKDNSRRVDNVLKLWIEARELPPK
KRYCYCEICLDMLYARTTSKPRASGDTVFVWGEHFEFNLPVAVRALRLHLRYRDSDKKR

KKDKAGYVGLVTPVATLAGRHFTEQWYPVTLPTGSGSGGMGSGGGSGSGSGGKG
KCCCBAYRIKBYOTMSILPMELYKEFAEYVTNHYRMLCAVLEBALNVKGKEEVASAL

VHILQSTGKAKDFLSDMAMSEVDFRMEHREHLIFRENTLATKAIEEYMRIGQKYLKDA

VFASWRLRCAERGEDRIADRLISASLFLRFLCPAIMSPSLFGLMQEYPDQTSRTLLT
IGEFIRALIESEENCEVFFIRCIASSLENGRNGRACGCEBGRVYVNRVETVYRNRN

DLGRELSTLHALLWEVLPQLSKEALLKGLPLRLLSDISTALRNPNIQRPSPQSER
IAKVIONLANFSKFTSKEDFLGFMNEFLELEWGSMMQFLIEISNEDLILINSSSESTI

RSQPMVLRGP^{SAEMQGYMMRD}LNSSIDLQSF^{MARGLNSSMDMAKLPSPINKEFFFF}
GGGKDLFYVSRPPLARSSPAYCTSSSDITEPEQKMLSVNKS^{VSMLDLQGDGPGGRLNS}

SSVSNLAAVGDLLHSSQASLTAAALGLRPAPAGRLSQSGSSITAAAGMRLSOMGVITDS
YBAOCIRPISEONPLFHMAADGPGPPAGHGSSGHGPPSSHHHHHHHHHHHRRGGEPPG

[illegible][illegible]

MPASERTVAWVSNNMPHLSADIESAHIETREEYKLKEYSKMDEE-
RLHMSNRKLEEYERRLLSQEEQTSKILMQYQARLEQSEKRLRQQVEKDSQIKS1103

```

PSLQADAGGGGAAPGPPRHG"
1043 a 1477 c 1314 g 967 t
PAGE COUNT

```

[illegible]

07 74: score 3502 8: DB 74: Length 4801:

	Query Match	87.2%	Score 5042.67	
	Best Local Similarity	99.3%	Pred. No. 0;	7. Indels 20; Gaps

Matches 3657; Conservative 0; Mismatches

Oy

169 atggaagcgcgcccgctggcccttcggccctcgcaaggcttccccgagc~33 33
|||||
|||

DB 4030 GCCCCCCCAGCCCCCAGCCCCCCCCACCCCGGTGC 4063

us-09-294-298-3.rqe

[illegible]

2329 aacaagagtgtgtccatgctggacctgcaggggcgacgggcctgggggacccc
|||||
3204

Db	4418	 GAGTACGAGCGGAGGCTGCTGCCAGGAAGACGACAGACAGATCTCGATCAGTAC	4477
Qy	3529	caagccgcctggagcagcagcagaagcgcctggagcagcagcagcagctggagaagactcc	3588
Db	4478	CNACCGCCTGGAGCAGACGACGAGAGCCCTTGAGCGACGACGAGGTGGAGAAGACTCC	4537
Qy	3589	cadatcaagagcatattgg-----caggctgatctggtggaggaggagct	3635
Db	4538	CAGATCAAGAGCATATTGGCAGCCCGTCCCTTCACGGCTGATGCTGGTGGAGGAGGAGCT	4597
Qy	3636	gcgccggagaccaccgcccatagctgagccgctgctcaaacccaagaagagctgcctcga	3695
Db	4598	GCGCCGGAGACCCCGCCATGCTGAGCGCGTGCTGAACCCNAGAAGAGGCTGCTCGA	4657
Qy	3696	cgctca-gagaggcagcttcccccttgggtccaaacaacccgcggtgtagcgtggcccc	3754
Db	4658	CGCTCAGGAGAGGACGCTTCCGCCCTTGGGTCCAAACAACCGCGCTGTGACGCTGGCCCC	4717
Qy	3755	acctggaaagcgctggccccccagccccaccccccccaccccccgctgcagatcacaga	3814
Db	4718	ACCTTGAAGCGGCTGGCCCCCCCCAGCCCCACCCCCCCCCCGGCTGCAGATCACAGA	4777
Qy	3815	gaacggcgagttccggaacacgcg	3838
Db	4778	GAACGGCGAGTTCGGAACACGCG	4801

RESULT	6					
LOCUS	HSDJ570F3	94770 bp	DNA	PRI	17-DEC-1999	
DEFINITION	Human DNA sequence from clone RP4-570F3 on chromosome 6. Contains a gene similar to Rattus norvegicus synaptic ras GTPase-activating protein p135, the C1CK0721Q.5 (polypeptide from patented cDNA Em:E06811) gene, the PHF1 (PHD finger protein 1) gene, the KNSL2 (kinesin-like 2) gene, the gene for acyl-protein thioesterase, (KSTs, STSS, GSXS and a CoG Island, complete sequence					

ACCESSION AL050332
 VERSION AL050332.15 GI:6010176
 KEYWORDS HTG; C1CK0721Q.5; CpG island; finger protein; GTPase-activating protein; kinesin-like; KNSL2; p135; PHD; PHF1; ribosomal protein L12; RPL12; SVN04P.

SOURCE
ORGANISM
Homo sapiens
Human.
LIZI, ADELIA, Sfrangini.

REFERENCE
Eukaryota: Metazoa; Craniata: Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 94770)
Mashrechi-Mohammadi, M.

AUTHORS

REFERENCE
AUTHORS
TITLE
JOURNAL
Submitted (17-DEC-1999) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
requests: clonerequest@sanger.ac.uk
Direct Submission
Mashreghi-Mohammadi, M.
1 (bases 1 to 94770)
mammalia; eutheria; feliformes; catarrhini; hominidae; homo.

COMMENT	REQUESTS
On Oct 4, 1999 this sequence version replaced gi:5870478. This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated human repeat sequence elements (e.g. Alu). Where the feature key is ambiguous, there is an annotation using the 'unsure' feature key.	gi:5870478 replaced by gi:5870479

This sequence was generated from part of bacterial clone contigs of human chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr6>

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:

Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at

Em:AI342398 Em:AA056366 Em:AI582796 Em:AI762297
Em:AI340198 Em:AI541387 Em:N34313 Em:AI567726 Em:AI364919
Em:AI310946 Em:AAJ181976 Em:AI042256 Em:R81592 Em:T49588
Em:AI1929571 Em:D52353 Em:AA035663 Em:N48534 Em:AI373814
Em:AA971499 Em:H28294 Em:AA703121 Em:AA446226 Em:AA379516
Em:AI313425 Em:W0287 Em:AA937820 Em:H21228 Em:AA471384
Em:N71594 Em:AA095718 Em:AA377858 Em:H25235 Em:H52756
/evidence=not_experimental
/product=dj570f3.2 (POLYPEPTIDE FROM PATENTED CDNA
EMBL:E06811)"

Query Match 22.3%; Score 920.8; DB 69; Length 94770;
Best Local Similarity 91.0%; Pred. No. 1.3e-167;
Matches 979; Conservative 0; Mismatches 97; Indels 0; Gaps 0;

QY 2150 acagctctatgacatgctcgctccctcccccacaaagagagaaccccccgccgcc 2209
DB 2744 ACAGCTCTATGACATGGCTCGGCTCCCTCCCAACCAAGAGAAAGCCACCCACACAC 2685
QY 2210 ctcccggtggggtaagacctgttctatgtgagcggccacacactggcccggtctccc 2269
DB 2684 CGCCTGGTGGTAAAGACCTGTCTATGTAAGCGGTCCACCCCTGGCCGTTCTCTCAC 2625
QY 2270 cagcatactgcagcagctcggacatcacagagccggagcagaagatgctgagtgtca 2329
DB 2624 CAGCATACTGCAGCAGCTCGGACATCATCAGAGCCAGACAGAGATGCTGAGTGTCA 2565
QY 2330 acaagagtgttcacatgctgacctcagagcagcagcggcctgagcgccttaacagca 2389
DB 2564 ACAAGAGTGTTCATGCTGGACATTACAGGGTGAATGGGCTGTGGCCGCTCAACAGCA 2505
QY 2390 gtagtgtttcaacctggcagctgtgtgggacctgttgcactcaagcagagcttcactga 2449
DB 2504 GCAGTGTTCGAACCTGCGCGCGCTAGGGGACCTGTGCTCAAGCAGGCGCTCGCTGA 2445
QY 2450 cagcagccttggggttgccgctcagcctcagcagcagcagcagcagcagcagcagcagc 2509
DB 2444 CAGCAGCCTTGGGGCTAGGGCTCGGCTCGCGACGCCCTCTCCAGGGAGTGGCTCAT 2385
QY 2510 ccatacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2569
DB 2384 CCATCAGCGGCTGGCATGCGCTCAGCCAGATGGGTGTGCACACAGACAGGTGTCCCTG 2325
QY 2570 ccagcaactgcgcatcctcttcttcccaagaacctctcttccatgatgctgctgctgctg 2629
DB 2324 CCCAGCACTGGCAATCCCGCTCTCTTCCAGAAACCCCTCTTCCACATGGCTGCTGATG 2265
QY 2630 gaccagggccccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2689
DB 2264 GGCCAGGTCCCCAGCGGCCATGAGGGGGCGGTGGCCATGGCCACCTTCTTCCCTCCATC 2205
QY 2690 acc 2749
DB 2204 ACCACCACCACCACCACCACCACCACCACCAGGTGGAGAGCCCGCTGGGACACCTTTGCC 2145
QY 2750 cgttccatggtatagcaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2809
DB 2144 CATTCATGGCTATAGCAAGAGTGGAGACCTCTCTTCCGGGGTCCCCAAGCCCCCTGCTG 2085
QY 2810 cctccatcttcacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2869
DB 2084 CTCTCATTCCTCATAGCCACAGCTACAGTGATGATGTTTGGACCCCTCTGGCACTGACTTCA 2025
QY 2870 cccgtcggcagctctcacttcaggacaacctacagcacatctctctcccccagagatca 2929
DB 2024 CCGCTGGCAGCTTTCATCTCCAGGACAACCTGCAGCACATGCTGTCCCTCTCCCAAGATCA 1965
QY 2930 ccatcggtcccccagagccagctccctcagggccagggaggggagggagggagggagggaggg 2989
DB 1964 CCATTGGTCCCCAGAGGCGGAGCCCCCTCAGGGCCCTGGAGGTGGGAGGGGCGGACCG 1905
QY 2990 gtgggggagcgtggggccagccacctctctgtcagaggggggcaaatctcagcagttgacag 3049

||||||| || |||||||||||| ||||| |||||||||||||||| |||||||||||||||| |||||||
Db 1904 GTGGGGTGGCGGGGCCAGCCCTCCATTCAGAGGGGCAAGTCTCAGCAGTTGACAG 1845
QY 3050 tgaagtgtcccaagaaccccccgctcagcggaacctatcgagtcctcccggaaccaa 3109
| || |||||||||||||| ||||| |||||||||||||||| |||||||||||||||| |||||||
Db 1844 TCAGCGCAGCCCAAGAAACCCCGCCATCCAGCGGGAATCTATTGCAGTCCCCAGACCAA 1785
QY 3110 gttatggtcctgcccctccacgagcaacagacagcctcagcaaaagagcgagcattg99ggca 3169
||||||| || |||||||||||||| ||||| |||||||||||||||| |||||||||||||||| |||||||
Db 1784 GTTATGGCCCGCCCGCTCCACGGCAACAGAGCTCTAGCAAGGAGGCGCAGCATTTGGGGGCA 1725
QY 3170 gcgggggacagcgtg9cg9agggggtg9ggggtcgaagccctccatcccaagcag 3225
| || |||||||||||||| ||||| |||||||||||||||| |||||||||||||||| |||||||
Db 1724 CGGGGGCAGCGGTGGCGGAGGGGTGGGGGCTGAAGCCCTCCATCACAAGCAG 1669
RESULT 7
AL161903/c AL161903 150956 bp DNA HTG 29-SEP-2000
LOCUS Homo sapiens chromosome 6 clone RP11-175A4, *** SEQUENCING IN
DEFINITION PROGRESS ***, 3 unordered pieces.
ACCESSION AL161903 GI:10039702
VERSION AL161903.12 HTG; HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 150956)
AUTHORS Tracey, A.
TITLE Direct Submission
JOURNAL Submitted (28-SEP-2000) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Sep 9, 2000 this sequence version replaced gi:9796876.
COMMENT ----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: BA175A4
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: Dye-terminator Big Dye; 99% of reads
Chemistry: Dye-primer-amersham; 0% of reads
Chemistry: Dye-primer Big Dye; 0% of reads
Consensus quality: 145680 bases at least Q40
Consensus quality: 147560 bases at least Q30
Consensus quality: 148600 bases at least Q20
Insert size: 150756; sum-of-contigs
Insert size: 148326; 10.8% error; agarose-fp
Quality coverage: 6.43x in Q20 bases; sum-of-contigs Quality
coverage: 6.54x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 54611: contig of 54611 bp in length
* 54612 54711: gap of 100 bp
* 54712 70293: contig of 15582 bp in length
* 70294 70393: gap of 100 bp
* 70394 150956: contig of 80563 bp in length.
Location/Qualifiers
1 .150956
/organism="Homo sapiens"
/db_xref="taxon:9606"

FEATURES
Source


```

Center: Sanger Genome Center
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: BA567N9
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 155424 bases at least Q40
Consensus quality: 160003 bases at least Q30
Consensus quality: 163252 bases at least Q20
Insert size: 167023; sum-of-contigs
Insert size: 186175; 5.7% error; agarose-fp
Quality coverage: 3.59x in Q20 bases; sum-of-contigs Quality
coverage: 3.22x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 24 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 5456: contig of 5456 bp in length
* 5457 5556: gap of 100 bp
* 5557 8767: contig of 3211 bp in length
* 8768 8867: gap of 100 bp

```

*	8868	17168:	contig of 8301 bp in length
*	17169	17268:	gap of 100 bp
*	17269	26591:	contig of 9323 bp in length
*	26592	26691:	gap of 100 bp
*	26692	31188:	contig of 4497 bp in length
*	31189	31288:	gap of 100 bp
*	31289	37579:	contig of 6291 bp in length
*	37580	37679:	gap of 100 bp in length
*	37680	45350:	contig of 7671 bp in length
*	45351	45450:	gap of 100 bp
*	45451	48354:	contig of 2904 bp in length
*	48355	48454:	gap of 100 bp
*	48455	63401:	contig of 14947 bp in length
*	63402	63501:	gap of 100 bp
*	63502	67724:	contig of 4223 bp in length
*	67725	67824:	gap of 100 bp in length
*	67825	70550:	contig of 2726 bp in length
*	70551	70650:	gap of 100 bp
*	70651	80864:	contig of 10014 bp in length
*	80665	80764:	gap of 100 bp
*	80765	83307:	contig of 2543 bp in length
*	83308	83407:	gap of 100 bp
*	83408	88383:	contig of 4976 bp in length
*	88384	88483:	gap of 100 bp
*	88484	93697:	contig of 5214 bp in length
*	93698	93797:	gap of 100 bp in length
*	93798	99354:	contig of 5557 bp in length
*	99355	99454:	gap of 100 bp
*	99455	110055:	contig of 11601 bp in length
*	110056	111155:	gap of 100 bp
*	111156	118645:	contig of 7490 bp in length
*	118646	118745:	gap of 100 bp
*	118746	122421:	contig of 3676 bp in length
*	122422	123231:	gap of 100 bp
*	123252	132468:	contig of 9947 bp in length
*	132469	132568:	gap of 100 bp
*	132569	137959:	contig of 5391 bp in length
*	137960	138059:	gap of 100 bp
*	138060	151691:	contig of 13632 bp in length
*	151692	151791:	gap of 100 bp
*	151792	158253:	contig of 6462 bp in length
*	158254	158353:	gap of 100 bp
*	158354	169323:	contig of 10970 bp in length

FEATURES SOURCE

[illegible]

misc_feature	48455. .63401 /note="assembly.fragment:00990 fragment_chain:2" 63502. .67724 /note="assembly.fragment:00045 fragment_chain:2" 67825. .70550 /note="assembly.fragment:00965 fragment_chain:2" 70651. .80664 /note="assembly.fragment:00637 fragment_chain:3" 80765. .83307 /note="assembly.fragment:01291 fragment_chain:3" 83408. .88383 /note="assembly.fragment:01317 fragment_chain:3" 88484. .93697 /note="assembly.fragment:01171 fragment_chain:4" 93798. .99354 /note="assembly.fragment:00288 fragment_chain:4" 99455. .111055 /note="assembly.fragment:00286 fragment_chain:4" 111156. .118645 /note="assembly.fragment:00303 fragment_chain:5" 118746. .122421 /note="assembly.fragment:01184 fragment_chain:5" 122522. .132468 /note="assembly.fragment:00501 fragment_chain:6" 132569. .137959 /note="assembly.fragment:01244 fragment_chain:6" 138060. .151691 /note="assembly.fragment:00032" 151792. .158253 /note="assembly.fragment:00779" 158354. .169323 /note="assembly.fragment:01270" /note="assembly.fragment:01269 g 42673 t 2375 others
misc_feature	41674 a 41332 c 41269 g 42673 t 2375 others

Query Match	22.3%	Score	920.8	DB	70	Length	169323
Best Local Similarity	91.0%	Pred. No. 1.le-167					
Matches	979	Conservative	0	Mismatches	97	Indels	0
Gaps	0						
Qy	2150	acaqctctatgacatgctcgcctccccctcccccaacgaagaaaccccgccgcgcc	2209				
Db	83616	ACASCTCTATGACATGCGCTCGCTCCCTCCCAACCAGGAAGACCAACCCCAACAC	83675				
Qy	2210	ctccccggtgggggtaaaagaccttctatgtgagcggcgccaccactgcccggctctccc	2269				
Db	83676	CGCCTGGTGGTGGTAAAGACCTCTTCATGTATGAGCGCTCCACCCTGGCCCGCTTCCTCAC	83735				
Qy	2270	cagcatactgcacgagcagctcgagacatcacagagccgagcagagaatgctgagtgctca	2329				
Db	83736	CASCATCTGCACGACGAGCTCGGACATCACAGAGCCACAGACAGAGATGCTGAGTGCA	83795				
Qy	2330	acaagaagtgtgtcaatgtgacactgagggcgacggcgctggggggcgcccttaacagca	2389				
Db	83796	ACAAGAGTGTGTCCATGTGTGACATTACAGGGTGATGGGCCCTGGTGGCCGCCCTCAACAGCA	83855				
Qy	2390	gtagtgttccaaacctggcagctgttggggacactgttgcaactcaagccaggcttcaactga	2449				
Db	83856	GCAGTCTTTTCGAACCTGCGGCCGTAGGGACCTGCTNGCACTCAAGCCAGGCGCTCGCTGA	83915				

gene	/organism="Homo sapiens" /db_xref="taxon:9606" /clone="as00087" /dev_stage="adult" /tissue_type="spleen" /note="vector:pBluescriptII SK plus" 1339..3363 /gene="FLJ000087"			
CDS	<1339..3363 /gene="FLJ000087" /note="For this clone, GeneMark analysis triggered an alert for spurious CDS split. The result of GeneMark analysis of the nucleotide sequence of this clone is accessible through http://www.kazusa.or.jp/NEDO. Start codon is not identified." /codon_start=1 /product="FLJ00087 protein" /protein_id="BAB15778.1" /db_xref="GI:10440490"			
KGPLPRAAAGAGVRAELWLDGALLARTAPRAGQGLFWHFHFEHFLPARRLSLR GLCPGSAFLCRVALAELEADAPAAAGLEKWFLLGAPAGALRARIRARRULVLP EYKELAEVTFHYARLCCALPEALPAQAKERLEAAVMRVLRATGRAOALVTDLGTAE LAPCGREALLPRENTLAPKATIDYMKVLVAODYLQETLGOVVRLECASTEDCEVDP VCASELPEHQARLNRCSREVFETIIHSDTWDFPABELGIVFSWREACKERGSEVGLPRL CPASLELRLCLAPSLFGLDPAPGAPPARTLIKAVIOLNANRAPFGKEAYM GFNMSLEERHPAMOCFLDOVAMVDVDAAPSGYGGSDGLALQALVHAQCLTFAE QWDTLELPPTLIRAIEQGVLVSPMRLPLPPAQVHSLSAGKEKPGFLAPDRLPK HTPLISKSSQSLARSRESWSARPEDEERLRRRPVQRTQSPVRRPARRQSGAGP PRPKSLMSGMPAPRPMPTWDSASLPKRPSPVQWQMDQPDQNRQALGTHRPVKNLAE LQCEVAALREOKVLSRLVESLSTOIRALTEQEBQLRGQLQDLDSRLRAG"				
785 a 1529 c 1261 g 712 t				
BASE COUNT				
ORIGIN				
Query Match 8.1%; Score 336.4; DB 34; Length 4287;				
Best Local Similarity 59.2%; Pred. No. 6.4e-55;				
Matches 574; Conservative 0; Mismatches 396; Indels 0; Gaps 0;				
QY	1010	ctgctgtgcgctaagccggttaccagacaatgagctatccctcccatgagctatata	1069	
Db	1769	CAGCGCTCGGGCGGGATTCGGCGCGTCCGCTGCGCGTGTGCGCTCCGAGCGCTACA	1828	
QY	1070	aggagttgcagaatatgtgaccaacaactaccgcacgtgtgtgctgctgtaggccg	1129	
Db	1829	AGGAGCTGGCGGAGTTCCTCACCTTCACCTATCGCGCGCCTTCGCGGGGCCCTGAGGCGCG	1888	
QY	1130	ccctcaatgtcaaggcaagagagtgctgtagtgcactggttcacatcctgcgaagca	1189	
Db	1889	CGCTCCTTGCAGGCCCAAGGAGAGTGGCGGCACGCCATGGTGGCGTGTCTGGCGGCCA	1948	
QY	1190	caggcaagcccaaggacttcctttcagacatgcccattgtcagaggtagacagggttcattg	1249	
Db	1949	CGGGCGGGCGCAGCGCGTGTGACTGACCTGGGCGACTCGCGAGCTGCGCGCTGTGGAG	2008	
QY	1250	agcgggaacactctatcttcggcgagacaacgctcccaactaaagccatagaagata	1309	
Db	2009	GCGGTGAGCGCTGCTGTTCCGGGAAACACATTGCCACCAAGGCTATCGATGAGTACA	2068	
QY	1310	tgagactgattgcccagaataatcctcaagatgccattggggaggttcataccgggctctgt	1369	
Db	2069	TGAAGCTCTGTGCACAGGATTTACTCTCAGAGAGACCTGGGACAGGTTGTGGCGGCTCTCT	2128	
QY	1370	atgaatctgagagaactgtgaagtagacccccatcaagtgcacagctccagctcgccag	1429	
Db	2129	GTGCTTCTACTGAGGACTGTGAAGTGACCCCAAGCAAAATGTCACGCTCGGAGCTGCAG	2188	
QY	1430	agcacaggccaaactgcgtagttgtctgagttgacctgtgcgaaggttggtcactcc	1489	
Db	2189	AGCACAGCGCCAGACTTTCGGAACAGCTTGCAGAGAGGCTCTTCGAAACCACTTATCCATCCT	2248	
QY	1490	attgcggtgttcccgaggagctgaaggaggtgttttgatcatctggcggtgcgctgtgcag	1549	

Qy	1660	ccagatgagcagacactcacaaacccctcacctctcatcgcaagggttatccagaacctggcc	1719
Db	24676	CGCTCGGACAGGGCACCCGAATCTCACACTGGTGGCCAGACCCCTGCAAACTTGGCC	24617
Qy	1720	aactttccaagatttacctcaaaaggagacttccagggtcatgaacagagttcttgag	1779
Db	24616	AATTTTACCCGCTTCCAGGGCAAGAGACTTTATGGAGTTTCTCAACGATTTCTCTCGAG	24557
Qy	1780	ctggagtggggtttctatcgagcaattcttgatgatatacaaacctgagcacactgac-	1838
Db	24556	CAGGAGGCGCTCGCATGTCGAACATGTTCTTGAGATTATATCAGCGCGCGGAGCACCCA	24497
Qy	1839	-----caacagcagcagttttgagggtacatagacttggccgcgagctctccaca	1890
Db	24496	GCTCCAGACTCGATCCTCGATTGGCCGGGTACATCGACCGGCAACAGATTGTCCATA	24437
Qy	1891	cttcaacgctctgtctctgggagggtgtgccccagctcagcaaggaaagccctctgaagctg	1950
Db	24436	CTACACAGTTTGTCTACGCGAAAGCCGTGGCCAAAGCTTCGCGAGGCCAGGACGACGACTG	24377
Qy	1951	ggcccgcttgccccggctcctcagcgacatcagcaagcgc	1989
Db	24376	GATCCGTTGCAGCATATCTTCGATGAAATCAGCCGAGCC	24338

RESULT 13				
AC012162				
LOCUS	AC012162	191504 bp	DNA	HTC
DEFINITION	Drosophila melanogaster chromosome X clone BACR01N10 (D1115) RPCI-98 01.N.10 map 16D-16F strain y; cn bw sp, *** SEQUENCING IN PROGRESS ***, 142 unordered pieces.			

ACCESSION	AC012162
VERSION	AC012162.9
KEYWORDS	HTG; HTGS_PHASE1.
SOURCE	fruit fly.
ORGANISM	Drosophila melanogaster
REFERENCE	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Eurygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephyridiodes; Drosophilidae; Drosophila. 1 (bases 1 to 191504).

REFERENCE	AUTHORS	TITLE
1 (bases 1 to 191504)		Sequencing of <i>Drosophila melanogaster</i>
	Butenikoff, S. E., Agabany, A., Arcana, T. T., Baxter, E., Blazej, R. G., Bubenik, C., Champe, M., Chavez, C., Chew, M., Ciesiolka, L., Doyle, C. M., Farfan, D. E., Galle, R., George, R. A., Harris, N. L., Hinkle, A., Hoskins, R. A., Houston, K. A., Hummasti, S. R., Karra, K., Kearney, L., Lee, B., Lewis, S., Li, P., Ling, H., Moshrefi, A. R., Moshrefi, M., Nixon, K., Pacleb, J. M., Park, S., Pfeiffer, B., Richards, S., Sethi, H., Swirskas, R. R., Wan, K. H., Webster, D., Woolley, P., Yang, S., Yee, M., Yu, C. and Rubin, G. M.	Unpublished

2 (bases 1 to 191504)
Unpublished

JOURNAL
REFERENCE
AUTHORS

Celinker, S. B., Agbayani, A., Arcaina, T. T., Baxter, E., Blazej, R. G.,
Butenhorff, C., Champe, M., Chavez, C., Chew, M., Ciesiolka, L.,
Doyle, C. M., Farfan, D. E., Galle, R., George, R. A., Harris, N. L.,
Hoskins, R. A., Houston, K. A., Hummasti, R., Karr, K., Kearney, L.,
Kim, E., Lee, B., Lewis, S., Li, P., Lomotan, M. A., Mazda, P.,
Moshrefi, A. R., Moshrefi, M., Nixon, K., Pacieb, J. M., Park, S.,
Pfeiffer, B., Poon, L., Sequeira, A., Sethi, H., Snir, E.,
Svirskas, R. R., Wan, K. H., Weinburg, T., Zhang, R., Zieran, L. L. and
Rubin, G. M.

TITLE
JOURNAL

COMMENT

Rubin, G.M.
Direct Submission
Submitted (21-OCT-1999) Drosophila Genome Center, Lawrence Berkeley Laboratory, MS 64-121, Berkeley, CA 94720, USA
On Feb 11, 2000 this sequence version replaced gi:6838825.
For further information about this sequence, including its location and relationship to other sequences, please visit our sequence archive Web site (<http://www.fruitfly.org/sequence/>) or send email to bdp@fruitfly.berkeley.edu. All contigs in this submission meet the following cutoffs: length >= 200 bases.
* NOTE: This is a 'working draft' sequence. It currently consists of 142 contigs. The true order of the pieces is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

*	1	456:	contig of 456 bp in length
*	457	536:	gap of unknown length
*	537	1004:	contig of 468 bp in length
*	1005	1084:	gap of unknown length
*	1085	1958:	contig of 874 bp in length
*	1959	2038:	gap of unknown length
*	2039	2248:	contig of 210 bp in length
*	2249	2328:	gap of unknown length
*	2329	2858:	contig of 530 bp in length
*	2859	2938:	gap of unknown length
*	2939	3325:	contig of 387 bp in length
*	3326	3405:	gap of unknown length
*	3406	4165:	contig of 760 bp in length
*	4166	4245:	gap of unknown length
*	4246	4716:	contig of 471 bp in length
*	4717	4796:	gap of unknown length
*	4797	5373:	contig of 577 bp in length
*	5374	5453:	gap of unknown length
*	5454	6050:	contig of 597 bp in length
*	6051	6130:	gap of unknown length
*	6131	6720:	contig of 590 bp in length
*	6721	6800:	gap of unknown length
*	6801	7461:	contig of 661 bp in length
*	7462	7541:	gap of unknown length
*	7542	8104:	contig of 563 bp in length
*	8105	8184:	gap of unknown length
*	8185	8505:	contig of 321 bp in length
*	8506	8585:	gap of unknown length
*	8586	9227:	contig of 642 bp in length
*	9228	9307:	gap of unknown length
*	9308	10091:	contig of 784 bp in length
*	10092	10171:	gap of unknown length
*	10172	11099:	contig of 928 bp in length
*	11100	11179:	gap of unknown length
*	11180	11747:	contig of 568 bp in length
*	11748	11827:	gap of unknown length
*	11828	12767:	contig of 940 bp in length
*	12768	12847:	gap of unknown length
*	12848	13398:	contig of 551 bp in length
*	13399	13478:	gap of unknown length
*	13479	14007:	contig of 529 bp in length
*	14008	14087:	gap of unknown length
*	14088	14867:	contig of 780 bp in length
*	14868	14947:	gap of unknown length
*	14948	15562:	contig of 615 bp in length
*	15563	15642:	gap of unknown length
*	15643	16859:	contig of 1217 bp in length
*	16860	16939:	gap of unknown length
*	16940	17986:	contig of 1047 bp in length
*	17987	18066:	gap of unknown length
*	18067	19341:	contig of 1275 bp in length
*	19342	19421:	gap of unknown length
*	19422	20926:	contig of 1505 bp in length
*	20927	21006:	gap of unknown length
*	21007	22153:	contig of 1147 bp in length
*	22154	22233:	gap of unknown length
*	22234	23236:	contig of 1003 bp in length
*	23237	23316:	gap of unknown length
*	23317	24607:	contig of 1291 bp in length
*	24608	24687:	gap of unknown length
*	24688	25620:	contig of 933 bp in length
*	25621	25700:	gap of unknown length
*	25701	26300:	contig of 600 bp in length
*	26301	26380:	gap of unknown length
*	26381	27488:	contig of 1108 bp in length
*	27489	27568:	gap of unknown length
*	27569	28215:	contig of 647 bp in length
*	28216	28295:	gap of unknown length

* 28296 contig of 1145 bp in length
 * 29441 gap of unknown length
 * 29520 contig of 1067 bp in length
 * 30587 gap of unknown length
 * 30667 gap of unknown length
 * 32042 contig of 1375 bp in length
 * 32123 gap of unknown length
 * 32123 contig of 1618 bp in length
 * 33741 gap of unknown length
 * 33821 contig of 1555 bp in length
 * 35376 gap of unknown length
 * 36258 contig of 803 bp in length
 * 36338 gap of unknown length
 * 36339 contig of 1665 bp in length
 * 38003 gap of unknown length
 * 38004 gap of unknown length
 * 38084 contig of 1426 bp in length
 * 39509 gap of unknown length
 * 39510 gap of unknown length
 * 39590 contig of 1130 bp in length
 * 40720 gap of unknown length
 * 40800 contig of 1654 bp in length
 * 42454 gap of unknown length
 * 42534 contig of 1308 bp in length
 * 43842 gap of unknown length
 * 43922 contig of 1489 bp in length
 * 45411 gap of unknown length
 * 45491 contig of 1292 bp in length
 * 46783 gap of unknown length
 * 46863 contig of 1708 bp in length
 * 48571 gap of unknown length
 * 48651 contig of 2392 bp in length
 * 51043 gap of unknown length
 * 51123 contig of 1638 bp in length
 * 52761 gap of unknown length
 * 52841 contig of 1872 bp in length
 * 54713 gap of unknown length
 * 54793 contig of 1729 bp in length
 * 56522 gap of unknown length
 * 56601 gap of unknown length
 * 56602 contig of 2337 bp in length
 * 58939 gap of unknown length
 * 59018 gap of unknown length
 * 59019 contig of 1819 bp in length
 * 60837 gap of unknown length
 * 60917 gap of unknown length
 * 62988 contig of 2071 bp in length
 * 63068 gap of unknown length
 * 63069 contig of 2515 bp in length
 * 65583 gap of unknown length
 * 65663 gap of unknown length
 * 65664 contig of 3111 bp in length
 * 68775 gap of unknown length
 * 68855 contig of 2381 bp in length
 * 71236 gap of unknown length
 * 71316 contig of 2798 bp in length
 * 74114 gap of unknown length
 * 74194 contig of 3212 bp in length
 * 77406 gap of unknown length
 * 77486 contig of 3451 bp in length
 * 80937 gap of unknown length
 * 81017 contig of 2341 bp in length
 * 83358 gap of unknown length
 * 83438 contig of 4019 bp in length
 * 87457 gap of unknown length
 * 87537 contig of 3996 bp in length
 * 91533 gap of unknown length
 * 91613 contig of 5531 bp in length
 * 91613 gap of unknown length
 * 97144 gap of unknown length
 * 97224 contig of 8370 bp in length
 * 105594 gap of unknown length
 * 105674 contig of 6398 bp in length
 * 112072 gap of unknown length
 * 112151 contig of 8860 bp in length
 * 121012 gap of unknown length
 * 121091 contig of 22022 bp in length
 * 121092 gap of unknown length
 * 143114 contig of 394 bp in length
 * 143588 gap of unknown length
 * 143667 gap of unknown length
 * 143668 contig of 260 bp in length

* 143928 144007: gap of unknown length
 * 144008 144842: contig of 835 bp in length
 * 144843 144922: gap of unknown length
 * 144923 145360: contig of 438 bp in length
 * 145361 145440: gap of unknown length
 * 145441 145915: contig of 475 bp in length
 * 145916 145995: gap of unknown length
 * 145996 146532: contig of 537 bp in length
 * 146533 146612: gap of unknown length

Query Match 5.8%; Score 241.8; DB 36; Length 191504;
 Best Local Similarity 50.6%; Pred. No. 3.6e-37;
 Matches 749; Conservative 0; Mismatches 682; Indels 48; Gaps 5;

QY 529 gaaaggcaaaatgattggaatctacagagggctgtgaaaccccaacaaaggacaacgc 588
 Db 113379 GAGCGGACCTTTGGATCTACTCGCTGCGCAAGTCGATCGCTCCGATGCGAGACACACG 113438

QY 589 cgcgggtagataacgctgctgaaactatggatcatatagaagctcgagagctgcccccaag 648
 Db 113439 CGTCGCACGACAACTCGCTGAAGATGTGGGTGTACGAGGCGAAAAATCTGCCGCCCAAG 113498

QY 649 aagcgalattactgcgagttatgcctggacgacatgctctatgcacgacacacttccaag 708
 Db 113499 AAGCGTTACTTTTGGGAACCTGCAATTGGACAAGACGCTGTACGGCCGCGACTTCGGTCAAG 113558

QY 709 cccgcctcagcctcaggagacactgtctttggggcgagcacttcgagtttaacaacctg 768
 Db 113559 CTGCAGAC-----GGATCTGCTGTTTGGGGGAGCAGCTTCGATTTTCCCGACATA 113609

QY 769 cctgtgtcgggctgctggctgcatctgtaccgtactcgagacaaagcgagaag 828
 Db 113610 CCGAGATTAAATGTATCACTACTTTAACGTTTTCGCTGAGGTGACAGAGAGAGAGCGG 113669

QY 829 gacaaggcaggctagctgtgctgtgactgtctccagtgccacccctggctggcgccac 888
 Db 113670 GACAAATACCAATTTGTGGGATCGGTGAAGATACCCGTCGACGATGTCACTCCAGATTG 113729

QY 889 ttacagagcagtggtacccctgacccctgccaacaggaagtggggctctgggggtatg 948
 Db 113730 CCTGCGAACTAGGTATCCCAT-----CTGAGCGACAGAGCGGGC 113771

QY 949 ggtctgggggaggggggtcagggggctcagggggcaaaaggagggctgt 1008
 Db 113772 GACAGCTCGGCAGGACATCGGGCGCGCGAGTGGTCAAGGACAAAGGACAAATTG 113831

QY 1009 cctgtgtcggctgaagcccgcttaaccagacaatagatctcctgcccattgagctatat 1068
 Db 113832 CCACGCTGAGGATCAAGGTGCTGTTTCCAGAGACACCGACATCTCCGCCCACTCAATGTGTAC 113891

QY 1069 aaggagtttgcagaaatgtgaccacacactaccgcatgtgtgtgtgctgtgagccc 1128
 Db 113892 GGCAACTTTTGGAGTACCTCAAGGAGAACTATAAGCGCGTGTGCGAGACCTCGAGCCG 113951

QY 1129 gcctcaatgtcaaggggcaaggagaggtcgctagtgcaactggttcaacatcctgcaaacg 1188
 Db 113952 GTATCGGAGTCAAGGCCAAGGAGGACATTTGGACAGGACCTGGTGTCTGCTGATCGACGCA 114011

QY 1189 acaggcaaggccaaggaactcctcttcagacatlggccaatgtcagaggtagaccggttcattg 1248
 Db 114012 CAGGATTGGCGGGCGCTTCTCTACCGATGTGGTGGCCCTCGATCTGCTGCGAGTTGGC 114071

QY 1249 gagggggaacacctatattcccgagagaacacgctcgccactaaagccaatagaagat 1308
 Db 114072 GATCAGAGG---CTTACGTTACAGGGCGCACTCTTTGGCCACCAAGAGATGAGGCGCATTC 114128

QY 1309 atgagactgattggccagaaaatcacctcaaggatgcatgctatggggagttcatccgggctctg 1368
 Db 114129 CTCAAGCTGACGGCGAAGACAGTATCTGCAGGACACACTATCCGACCACTAAACAGACCTA 114188

QY 1369 tatgaatctgaggagaaactgtgaagttagaccccccaatgaagtga-----cagcgctcc 1419
 Db 114189 ATTCAGTCGAGAGGAGTGTGCGAGGTGGATCCCAAGAGAGCGGCTTCGTGCGCGGGT 114248

Qy 1420 agtctgacagaccagcgaacccctgcgagtgctgtgagtgccctgtgcaagtg 1479
 Db 114249 TCGCTGACGACAGCAGGCGCGCTTGCCTGGCGCGCTCCGAGGGCGTGGCGATC 114308
 Qy 1480 gtcaactccattgctgtcccgagggagctgaagaggtgtttgcatcatcgogctg 1539
 Db 114309 TTCAATCGCAGCAAGCAATTTTCGCCGCCAGTTACGAAATGCTTTCGCGAGTTCGGGAG 114368
 Qy 1540 cgtgtgacagagcgggcgagagacattgctgacagctgacagcgcctcgtcttc 1599
 Db 114369 CGCTTTCGACAGCGTGGCGCGCTCAGGATATGCGGCAACCTGATCTCGGCGAGCAATTTTC 114428
 Qy 1600 ctgcgtctctctgcgcgcgcacatgctgcgcagctctgttggactgatgcagagtagc 1659
 Db 114429 CTGCGCTTCCTGTGCGCGGCATCTCTGCGCGTCTCTTCAATATCACCAGCGAATG 114488
 Qy 1660 ccaatgacagacccctacgaacccctcaacccctatcgccaaaggttatccagaaacctgccc 1719
 Db 114489 CCGTCCGACAGGCGCCACCGCAATCTCACACTGCTGTCGCGCAAGACCTGCAAAACATGTGCC 114548
 Qy 1720 aactttcccaagtttaactcaagaggagactctctctggtgcttcatgaacagagtttctgag 1779
 Db 114549 AATTTCACCCGCTCCAGGGCAAGAGAACTTTATGAGTCTTCTCAACGATTTCTCTCGAG 114608
 Qy 1780 ctgagtggttctatgacgaattctgtatgagatatcaaacctggacacactgac- 1838
 Db 114609 CAGGAGCCGCTCGATGCAACAGTCTTCTGGAGATTATATCCACGCGCGGAGCACCACA 114668
 Qy 1839 -----caacagcagcagttttgagggctacatagacttggggcgagctctccaca 1890
 Db 114669 GCTCCAGACTCGATCTCTCGATTGCGCGGCTATCATCGACGAGGCAAAAGTTGTCCATA 114728
 Qy 1891 ctcaagcctctctgagagtgctgcccagctcagcaagagcctcctcctgagctg 1950
 Db 114729 CTACACAGTTGCTACGCAAAACCTTGGCCAGCTGCCGGAGGCGCAGGACGAGCTG 114788
 Qy 1951 ggcgcgtgccccgctctcagcagcatcagcacagcc 1989
 Db 114789 GATCCGTTGCAGCATATCTCGATGAATCAGCGGAGCC 114827

RESULT 14
 AC012161/c
 LOCUS
 DEFINITION Drosophila melanogaster chromosome X clone BACR01122 (D1114)
 RPII-98 01.1.22 map 16B-16D strain y; cn bw sp. *** SEQUENCING IN
 PROGRESS ***, 142 unordered pieces.
 AC012161
 AC012161.7 GI:6435885
 HTG; HTGS_PHASE1.
 SOURCE fruit fly.
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 202741)
 Ceiniker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,
 Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,
 Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,
 Hinkle,A., Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K.,
 Kearney,L., Lee,B., Lewis,S., Li,P., Ling,H., Moshrefi,A.R.,
 Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S., Pfeiffer,B.,
 Richards,S., Sethi,H., Svirskas,R.R., Wan,K.H., Webster,D.,
 Woolley,P., Yang,S., Yee,M., Yu,C. and Rubin,G.M.
 Sequencing of Drosophila melanogaster
 Unpublished
 2 (bases 1 to 202741)
 Ceiniker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,
 Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,
 Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,
 Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L.,
 Kim,E., Lee,B., Lewis,S., Li,P., Lomotan,M.A., Mazda,P.,

Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S., Pfeiffer,B., Poon,L., Sequeira,A., Sethi,H., Snit,E., Svirskas,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and Rubin,G.M.
 Direct Submission
 Submitted (21-OCT-1999) Drosophila Genome Center, Lawrence Berkeley Laboratory, MS 64-121, Berkeley, CA 94720, USA
 On Nov 16, 1999 this sequence version replaced gi:6425635.
 For further information about this sequence, including its location and relationship to other sequences, please visit our sequence archive web site (http://www.fruitfly.org/sequence/) or send email to bdgs@fruitfly.berkeley.edu. All contigs in this submission meet the following cutoffs: length >= 200 bases.
 * NOTE: This is a 'working draft' sequence. It currently consists of 142 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 509: contig of 509 bp in length
 510 589: gap of unknown length
 590 1201: contig of 612 bp in length
 1202 1281: gap of unknown length
 1282 2394: contig of 1113 bp in length
 2395 2474: gap of unknown length
 2475 3512: contig of 1038 bp in length
 3513 3592: gap of unknown length
 3593 4500: contig of 908 bp in length
 4501 4580: gap of unknown length
 4581 5465: contig of 885 bp in length
 5466 5545: gap of unknown length
 5546 6253: contig of 708 bp in length
 6254 6334: gap of unknown length
 6334 7105: contig of 772 bp in length
 7106 7186: gap of unknown length
 7187 8077: contig of 892 bp in length
 8078 8157: gap of unknown length
 8158 9047: contig of 890 bp in length
 9048 9127: gap of unknown length
 9128 9530: contig of 403 bp in length
 9531 9610: gap of unknown length
 9611 10173: contig of 563 bp in length
 10174 10253: gap of unknown length
 10254 10841: contig of 588 bp in length
 10842 10921: gap of unknown length
 10922 11807: contig of 886 bp in length
 11808 11887: gap of unknown length
 11888 13165: contig of 1278 bp in length
 13166 13245: gap of unknown length
 13246 14526: contig of 1281 bp in length
 14527 14606: gap of unknown length
 14607 15801: contig of 1195 bp in length
 15802 15881: gap of unknown length
 15882 16640: contig of 759 bp in length
 16641 16720: gap of unknown length
 16721 17679: contig of 959 bp in length
 17680 17759: gap of unknown length
 17760 18797: contig of 1038 bp in length
 18798 18877: gap of unknown length
 18878 19955: contig of 1078 bp in length
 19956 20035: gap of unknown length
 20036 20931: contig of 896 bp in length
 20932 21011: gap of unknown length
 21012 21643: contig of 632 bp in length
 21644 21723: gap of unknown length
 21724 23416: contig of 1693 bp in length
 23417 23496: gap of unknown length
 23497 24457: contig of 961 bp in length
 24458 24537: gap of unknown length
 24538 25955: contig of 1418 bp in length
 25956 26035: gap of unknown length

TITLE
JOURNAL

COMMENT

* 26036 27838: contig of 1803 bp in length
* 27839 27918: gap of unknown length
* 27919 29686: contig of 1768 bp in length
* 29687 29766: gap of unknown length
* 29767 30976: contig of 1210 bp in length
* 30977 31056: gap of unknown length
* 31057 32460: contig of 1404 bp in length
* 32461 32540: gap of unknown length
* 32541 34644: contig of 2104 bp in length
* 34645 34724: gap of unknown length
* 34725 36341: contig of 1617 bp in length
* 36342 36421: gap of unknown length
* 36422 37864: contig of 1443 bp in length
* 37865 37944: gap of unknown length
* 37945 39424: contig of 1480 bp in length
* 39425 39504: gap of unknown length
* 39505 41341: contig of 1837 bp in length
* 41342 41421: gap of unknown length
* 41422 42970: contig of 1549 bp in length
* 42971 43050: gap of unknown length
* 43051 44419: contig of 1369 bp in length
* 44420 44499: gap of unknown length
* 44500 46484: contig of 1985 bp in length
* 46485 46564: gap of unknown length
* 46565 48446: contig of 1882 bp in length
* 48447 48526: gap of unknown length
* 48527 50693: contig of 2167 bp in length
* 50694 50773: gap of unknown length
* 50774 53578: contig of 2805 bp in length
* 53579 53658: gap of unknown length
* 53659 55753: contig of 2095 bp in length
* 55754 55833: gap of unknown length
* 55834 58457: contig of 2624 bp in length
* 58458 58537: gap of unknown length
* 58538 60171: contig of 1634 bp in length
* 60172 60251: gap of unknown length
* 60252 63621: contig of 3370 bp in length
* 63622 63701: gap of unknown length
* 63702 66219: contig of 2518 bp in length
* 66220 66299: gap of unknown length
* 66300 69074: contig of 2775 bp in length
* 69075 69154: gap of unknown length
* 69155 72903: contig of 3751 bp in length
* 72906 72985: gap of unknown length
* 72986 75963: contig of 2978 bp in length
* 75964 76043: gap of unknown length
* 76044 78866: contig of 2823 bp in length
* 78867 78946: gap of unknown length
* 78947 82045: contig of 3099 bp in length
* 82046 82125: gap of unknown length
* 82126 85135: contig of 3010 bp in length
* 85136 85215: gap of unknown length
* 85216 88550: contig of 3335 bp in length
* 88551 88630: gap of unknown length
* 88631 92611: contig of 3981 bp in length
* 92612 92691: gap of unknown length
* 92692 98709: contig of 6018 bp in length
* 98710 98789: gap of unknown length
* 98790 104148: contig of 5358 bp in length
* 104149 104227: gap of unknown length
* 104228 111796: contig of 7569 bp in length
* 111797 111876: gap of unknown length
* 111877 118488: contig of 6612 bp in length
* 118489 118568: gap of unknown length
* 118569 126924: contig of 8356 bp in length
* 126925 127004: gap of unknown length
* 127005 137051: contig of 10047 bp in length
* 137052 137131: gap of unknown length
* 137132 148859: contig of 11728 bp in length
* 148860 148939: gap of unknown length
* 148940 149554: contig of 615 bp in length
* 149555 150197: contig of 563 bp in length
* 149635

* 150198 150277: gap of unknown length
* 150278 150907: contig of 630 bp in length
* 150908 150987: gap of unknown length
* 150988 151646: contig of 659 bp in length
* 151647 151726: gap of unknown length
* 151727 152074: contig of 348 bp in length
* 152075 152154: gap of unknown length
* 152155 152902: contig of 748 bp in length
* 152903 152982: gap of unknown length
* 152983 153624: contig of 642 bp in length
* 153625 153704: gap of unknown length
* 153705 154242: contig of 538 bp in length
* 154243 154322: gap of unknown length
* 154323 154739: contig of 417 bp in length
* 154740 154819: gap of unknown length
* 154820 155338: contig of 519 bp in length
* 155339 155418: gap of unknown length
* 155419 155993: contig of 575 bp in length
* 155994 156073: gap of unknown length
* 156074 156806: contig of 733 bp in length
* 156807 156886: gap of unknown length
* 156887 157541: contig of 655 bp in length
* 157542 157621: gap of unknown length
* 157622 158208: contig of 587 bp in length
* 158209 158288: gap of unknown length

Query Match 5.8%; Score 241.8; DB 36; Length 202741;
Best Local Similarity 50.6%; Pred. NO. 3.5e-37;
Matches 749; Conservative 0; Mismatches 682; Indels 48; Gaps 5;

Qy 529 gaaaggacaaatgattgaaatcacagagggtgtgaaacccaacaaagcaacagc 588
Db 126449 GACGGCAGCTTTGGATCTACTCGCTGCGCAAGTCGATCGCTCGCAATGACAGACACAG 126390

Qy 589 cgcgggttagataacgtcgaaactatggatcatagaagctcgagagctgcccccaag 648
Db 126389 CGTCGACGACACATCGCTGAAGATGCGGTGTACGAGGCGGAAATCTCCGCCCAAG 126330

Qy 649 aacgatattactcaggtatgctcgacacacatgctcgcacgacacacacttcaag 708
Db 126329 AACGGTTACTTTTGGCAACTGCANTTGGACAAGACGCTGTACGGCGGACCTTCGGTGAG 126270

Qy 709 ccccgctcagcctcagagacacactgtcttttggggcgagcacttcgagtttaacacctg 768
Db 126269 CTCGACAC-----GGATCTGCTGTTTGGGGGAGACATTCGATTTCCCGACATA 126219

Qy 769 cctgctgtccggcgctcggtgcatctgtacctgtactcggacacaaaggaagaag 828
Db 126218 CCCGAGATTAAATGTATCACTGTTAACGTTTCCGTGAGGTGGACAGAAAGAACGG 126159

Qy 829 gacaaggcaggttaactgtgctgtgactgttccagtggtccacctggctggcgccac 888
Db 126158 GACAAATACCAATTTGTGGATCGGTGAAGATACCCGTGCACGATGTACCTCCAGATTG 126099

Qy 889 ttacagagcagtggttaacccgtgacctgacccctgccaacaggaagtgggggtatg 948
Db 126098 CCCTGCGAACATGGTATCCCATATA-----CTGAGCGCAACAGCGCGGC 126057

Qy 949 ggcctcggggagaggggtcaggggcggtcagggggcaaggggaaaggaggtcgt 1008
Db 126056 GACAGTCTGGCAGACATCGGGCGGCGCGCAGTGGGTCAAGGACAGGACCAATTG 125997

Qy 1009 cctgctgtcggctgaagcccggttacacagacaatgagtatcctgcctcgaagactalat 1068
Db 125996 CCCACGCTGAGGATCAAGTGTCTGTTCCAGAGCACCGACCTGCCCATCAATGTGTAC 125937

Qy 1069 aaggagtttgcaaatatgtgacccaacactaccgcatgtgtgtgcgctgctgagccc 1128
Db 125936 GGCAACTTTTGTGACGTACCTCAAGGAGAACTATAAGCGCGGTGCGAGACCTCGAGCTG 125877

Qy 1129 gccctcaatgtcaaggcgcaaggaggtcgctagtgcactggttccactcctcctcacaagc 1188
Db 125876 GTGATCGGAGTCAAGGCCAAGGAGGACATTTGACAGGACACTGGTGTCTGTATGACGCA 125817


```

YIAESGRPIGEKFMWFCFTSISGAVTALVILMSLWEKFQTNPTITGLDTHFNQNVVFP
ITVQVPEAAEDHDKTYEKVNTLANDEAQAQMYTPFLRIILTSNFENVRDKAVLSQS
IPQMLLDAFTIRWEAFEGHIDCKNVFVCSQYRDEIDPCDHFIETHTYHGRFCAFYNSR
FKSPTEDVKTGAPHDLXYETDKKWLFFIPNFSKIFIFSNEEFYFQDFNAQIDWSEP
OLVBRISKNTVTTDDARQLSIGORAKCIFSEVKLVNFPDQAYTFSSCKMCKRNKAI
KLCKNPFPKPIPNVPMCSIKDFDLDEFKSNLTINIKDCIQCLSCSKVFNIDKLI
KMSDRPSGLVIEFTWPIIRYKREVLFGWVDLLVSEGGIASIPLGFSILSGVEIILY
YFTLRACWYKRWQELYEIEKIRQEPPEPKIDLKLSKSNPRIGDPTISAVLKVP
"
AEEAGNPJGGYORRHEKDYSLNTKNTYKTPKVPVDYGYTSKTSNWTIYDHSQYMP
"
complement(join(21273..21792,21868..21999,22058..22191,
22257..22823,22889..23887,24168..24320,24398..25116,
25191..25449))
/sequence="CG8465"
/product="CT24743"
/db_xref="FLYBASE:FBan0008465"
/db_xref="FLYBASE:FBgn0030845"
complement(join(21273..21792,21868..21999,22058..22191,
22257..22823,22889..23887,24168..24320,24398..25116,
25191..25370,27861..28246))
/sequence="CG8465"
/product="CT24757"
/db_xref="FLYBASE:FBan0008465"
/db_xref="FLYBASE:FBgn0030845"
complement(join(21273..21792,21868..21999,22058..22191,
22257..22823,22889..23887,24168..24320,24398..25116,
25191..25370,27861..28115,28453..28555))
/sequence="CG8465"
/product="CT24733"
/db_xref="FLYBASE:FBan0008465"
/db_xref="FLYBASE:FBgn0030845"
complement(<21273..>28555)
/sequence="CG8465"
/map="16B4-16B7"
/db_xref="FLYBASE:FBan0008465"
/db_xref="FLYBASE:FBgn0030845"
complement(join(21399..21792,21868..21999,22058..22191,
22257..22823,22889..23887,24168..24320,24398..25116,
25191..25370,27861..28092))
/sequence="CG8465"
/sequence="CG8465 gene product [alt 1]"
/codon_start=1
/db_xref="FLYBASE:FBan0008465"
/db_xref="FLYBASE:FBgn0030845"
/protein_id="AAF48735.1"
/db_xref="GI:7293357"
/translation="MSTYFGVIPTSKAGCFEGSVSQICGSIATAVNIKPSNPASGSAS
VASGSPGSAASVOTGNADGSAATKYEDPDYPPDSPMLIFTEKSKALDILRHVKEA
RLREFPNLEAESYVQGFESIEALKRFCKAKPEKPIPIISGSGYKSPSTSDNSCS
SSPTNGSGFTIPLGSSNSNLLSDSPSSPSSSNVIANGROOQOQOQOQOQOQOQ
PDVSGEGPPFPAPTKQELVEFRKQIEGGHIDRVKRIWENPRFLISSGDTPTSLKEG
RYNAMHCAOVNARIAIOLLKTIISDREFTLQYVGKSGKMCALNLSLDYLNMPDK
DKRGEPFLHFAKNGHVAVELVSVPECKSLRNHGEKPEKIEICLRNANATHVTLK
KLELLLDHPVPLVRSQNTLPPKVGOPSPKDPNLOHKAADDYEGISVDLATSALA
GPMREKAMNFRWKTPPRVSNVNSPPLAGSPFVKVPTSPLEFPATPIRMKPDPLF
SGRRVLFSPAELAEATSPKPNVNGTNECHNNNNKVPVPLEFPATPIRMKPDPLF
MAYRNNSFDSPLADDSQILDSLSRLNSLDSFRERHKNIDTIEKGLVGVGRQL
ARQOLEWREYVDFLSDIDGTEGLARLEAYFLKTEQQAADKSETVMNFVFAHLHGY
DSMAGEQQQLRKDKNEAGATSPSAGVMTPTCVKSLQVFAKRTITLKNIGNMV
SINDTLCELRKRLKSLIVSFKNPVTGNECHNNNNKVPVPLEFPATPIRMKPDPLF
OLLOMLSLRQLLADDERGHLGCVCASILLMLEQATSAVHLPTDLKTEIGCCAAW
ETEOCCALVDANLSRKTSSRKRTSLRAAAVVSQGLQDTSGSTSGSALHASLGVG
PPOSLSGISMOLPEPVEYIFGNEPTKRDLDVNLNSVDIDKETLPHVYAWKTAME
STSLGASRVASAKDARQQDDSDYEDQVIFDCTNVTLPYSGSSEDEENFT
SYSCAEMNLNVKQKPEPYSGTSSSHNSQPLHPRKRLATPKLNAVVSRRGSGPLT
APVTPRLARTPSAASIOVASETNGESVGTAVTASPILSFAALTAATQSFQPLNKR
GLFSQYRDORSYNEGDTPLGNRN"
complement(join(21399..21792,21868..21999,22058..22191,
22257..22823,22889..23887,24168..24320,24398..25116,
25191..25410))
/sequence="CG8465"
/sequence="CG8465 gene product [alt 3]"
/codon_start=1
/db_xref="FLYBASE:FBan0008465"
/db_xref="FLYBASE:FBgn0030845"
/protein_id="AAF48737.1"
/db_xref="GI:7293359"
/translation="MPTHQCHRHDAADSPMLIFTEKSKALDILRHVKEARLREFP
NLEAESYVQGFESIEALKRFCKAKPEKPIPIISGSGYKSPSTSDNSCS
SSPTNGSGFTIPLGSSNSNLLSDSPSSPSSSNVIANGROOQOQOQOQOQOQOQ
GPPFPAPTKQELVEFRKQIEGGHIDRVKRIWENPRFLISSGDTPTSLKEGCRYNAMH
CAOVNARIAIOLLKTIISDREFTLQYVGKSGKMCALNLSLDYLNMPDKGRGE
TPLHFAKNGHVAVELVSVPECKSLRNHGEKPEKIEICLRNANATHVTLKLELL

```

```

Query Match      5.8%; Score 241.8; DB 28; Length 300994;
Best Local Similarity 50.6%; Pred. No. 3.2e-37;
Matches 749; Conservative 0; Mismatches 682; Indels 48; Gaps 5;

Qy 529 gaaaggacaaatgattgagaatctacagagggtgtgaaacccaacagcaacagc 588
Db 196157 GAGCGGACCTTTGGATCTACTCGCTGCCAAGTCGATCGCTCCGATGCGAGACACG 196216

Qy 589 cgcgggtagataaacgtgctgaaacttgatcatagaagctcgagagctgcccccaag 648
Db 196217 CGTCGACGACAACTCGCTGAAGATGTGGGTGTACGAGCGGAAAAATCTGCCGCCAAG 196276

Qy 649 aagcgatatctgcagagtatgcctggagacagatctctatgcagcaactccaag 708
Db 196277 AAGCGTTACTTTTGGCAACTGCAATTGGACAAGACGCTGTACGCGCGGACTTCGGTGAAG 196336

Qy 709 ccccgctcagcctcagagacagactgtcttttggggcgagcacttcgagtttaacaaactg 768
Db 196337 CTGCAGAC-----GGATCTGCTGTTTGGGGGAGACATTCGATTTCCCGACATA 196387

Qy 769 cctgctgctcgggcgctcgggctgcatctgtaccctgactcgagcaaaagcgaagaag 828
Db 196388 CCGGAGATTAAATGTGATCACTGTTTAACTGTTTCCGTGAGGTGGACAAAGAAAGACGG 196447

Qy 829 gaaacagcaggtacgttgctgctggtgactgttccagtgagccctggtggtggtcgcac 888
Db 196448 GACAAATACCAATTTCTGGGATCGGTGAAGATACCCGTGCAGGATGTACCTCAGATTG 196507

Qy 889 ttacagagcaggtggttacccctgaccctgccaacaggaagtgggggtctctgggggtatg 948
Db 196508 CCCTGCGCAACAATGGTATCCCATATA-----CTGAGCGCAAGCGCGGC 196549

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 18, 2001, 11:15:27 ; Search time 231 Seconds
(without alignments)
2884.142 Million cell updates/sec

Title: US-09-294-298-3
Perfect score: 4134
Sequence: 1 taaggccccccaccgacc.....tgggggtatcccccatct 4134

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 280836 seqs, 80580151 residues
Total number of hits satisfying chosen parameters: 561672

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents.NA.*
1: /cgn2.6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2.6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2.6/ptodata/2/ina/6_COMB.seq.*
4: /cgn2.6/ptodata/2/ina/PCFUS_COMB.seq.*
5: /cgn2.6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	63.6	1.5	3148	3	US-08-909-954-1
2	63.4	1.5	7218	1	US-08-232-463-14
3	62	1.5	3117	3	US-08-908-954-3
4	61	1.5	2580	3	US-09-050-863-2
5	61	1.5	5452	2	US-09-130-114-1
6	61	1.5	10596	1	US-07-884-811-15
7	61	1.5	10596	1	US-07-885-971-15
8	61	1.5	10596	1	US-08-087-783A-15
9	61	1.5	10596	1	US-08-194-088B-15
10	61	1.5	10596	2	US-08-194-087-15
11	61	1.5	10596	4	PCF-US93-04648-15
12	57.6	1.4	2338	1	US-08-425-069-1
13	57.6	1.4	2338	2	US-08-317-844B-1
14	56	1.4	16442	3	US-08-781-891-208
15	54.6	1.3	3489	2	US-08-728-323A-1
16	54.6	1.3	32207	2	US-08-770-379-20
17	54.4	1.3	9551	2	US-08-056-200-93
18	54.4	1.3	9551	2	US-08-800-644-93
19	54	1.3	543	5	5273901-6
20	53.6	1.3	1995	1	US-08-425-069-3
21	53.6	1.3	1995	2	US-08-317-844B-3
22	53	1.3	1137	5	5171843-8
23	52.2	1.3	1847	3	US-08-675-885-4
24	51.4	1.2	657	5	5171843-10
25	51.4	1.2	152331	3	US-09-128-155-16
26	51.2	1.2	2793	1	US-08-209-747-1
27	51.2	1.2	2793	1	US-08-458-298-1
28	51.2	1.2	51259	3	US-08-781-891-209

29	49	1.2	2040	2	US-08-533-669A-5	Sequence 5, Appli
30	47.8	1.2	3337	1	US-08-072-610-1	Sequence 1, Appli
31	47.8	1.2	3337	2	US-08-719-822B-1	Sequence 1, Appli
32	47.6	1.2	397	3	US-09-253-691-3	Sequence 3, Appli
33	46	1.1	234	1	US-08-469-802B-3	Sequence 3, Appli
34	46	1.1	234	2	US-08-267-803B-3	Sequence 1, Appli
35	46	1.1	2818	1	US-08-366-276-1	Sequence 14, Appli
36	45	1.1	7218	1	US-08-232-463-14	Sequence 24, Appli
37	45	1.1	3456	1	US-08-190-687B-24	Sequence 7, Appli
38	45	1.1	4307	1	US-08-190-687B-7	Sequence 7, Appli
39	44.8	1.1	44377	2	US-08-804-227C-7	Sequence 7, Appli
40	44.8	1.1	44377	2	US-08-804-198-1	Sequence 1, Appli
41	44.6	1.1	533	5	5482709-5	Patent No. 5482709
42	44.4	1.1	1599	3	US-08-853-733B-1	Sequence 1, Appli
43	44.4	1.1	1603	3	US-08-675-885-6	Sequence 6, Appli
44	44.4	1.1	1931	2	US-09-130-114-2	Sequence 2, Appli
45	43.8	1.1	2385	1	US-08-393-333-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-08-909-954-1
; Sequence 1, Application US/08909954A
; Patent No. 6100058
; GENERAL INFORMATION:
; APPLICANT: Allen, Maxine J.
; APPLICANT: Buckler, Alan J.
; TITLE OF INVENTION: GAP12 Genes and their Uses
; FILE REFERENCE: SEQ-11P
; CURRENT APPLICATION NUMBER: US/08/909,954A
; CURRENT FILING DATE: 1997-08-12
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3148
; TYPE: DNA
; ORGANISM: H. sapiens
US-08-909-954-1

Query Match 1.5%; Score 63.6; DB 3; Length 3148;
Best Local Similarity 50.0%; Pred. No. 7.4e-06;
Matches 159; Conservative 0; Mismatches 159; Indels 0; Gaps 0;

QY	1414	gcgtccagctgagcagaccagccacacgtgctgagtgagtgccctgtgc	1473
DB	1518	gagaccagctggggtgctgacgggtctacctggggcccatcgtggacccatctg	1577
QY	1474	aaggtggtaactcccatcctgctgttcccgaggaggtgaagaggtgtttgcatc	1533
DB	1578	tcctggggcgctgcccgcccatcgctcctcaagcagctgacccggcgagtg	1637
QY	1534	cggctgcgctgagagcgggggcgagacatgctgacagctgacagccctgc	1593
DB	1638	gaggagcgcttcccccggcgccagcagcaggtgtgaaagctacgtgaggtatt	1697
QY	1594	ctcttcctgctctctctctcctgcccggccatcgtgccagtgctgttgactgat	1653
DB	1698	ctcttcctgctctctcctgcccggccatcgtgccagtgctgttgactgat	1757
QY	1654	gagttaccagatgagcagcctcagacacccctcaccctcagcaggtttaccag	1713
DB	1758	caacacgaggaccccccagcagtagcgcctcactgctgtgctgccaaggtgtg	1817
QY	1714	ctggccaaacttttccaaag	1731
DB	1818	attggaacacctggggccag	1835

RESULT 2
US-08-232-463-14/c


```

; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232.463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935.313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZ9pt-Fls
; US-08-232-463-14

```

```

Query Match 1.5%; Score 63.4; DB 1; Length 7218;
Best Local Similarity 3.3%; Pred. No. 1.1e-05;
Matches 13; Conservative 233; Mismatches 149; Indels 0; Gaps 0;

Qy 3317 acatcgagagtcacacattgacgaggaagagtagacagctgaagagtagctacgaagtcca 3376
Db 1436 ACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1377

Qy 3377 tggacgagagccgactgacaggggtggaagaggtacgagagtagacacactcactgaag 3436
Db 1376 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1317

Qy 3437 aaagctacacatgtccacaggaagctggaagagtagcagcagcagcagcagcagcagcag 3496
Db 1316 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1257

Qy 3497 aaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 3556
Db 1256 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1197

Qy 3557 gcttgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 3616
Db 1196 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1137

Qy 3617 tgcgtgagagagagctgcgcgggagcagcagcagcagcagcagcagcagcagcagcagcagcag 3676

```

```

Db 1136 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1077
Qy 3677 ccaagaagagctgctgcagcgtctcagagagcagc 3711
Db 1076 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1042

RESULT 3
US-08-909-954-3
; Sequence 3, Application US/08909954A
; Patent No. 6100058
; GENERAL INFORMATION:
; APPLICANT: Allen, Maxine J.
; APPLICANT: Buckler, Alan J.
; TITLE OF INVENTION: GAP12 Genes and their Uses
; FILE REFERENCE: SEQ-11P
; CURRENT APPLICATION NUMBER: US/08/909,954A
; CURRENT FILING DATE: 1997-08-12
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 3117
; TYPE: DNA
; ORGANISM: M. musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (0)...(0)
; US-08-909-954-3

```

```

Query Match 1.5%; Score 62; DB 3; Length 3117;
Best Local Similarity 49.7%; Pred. No. 1.7e-05;
Matches 158; Conservative 0; Mismatches 160; Indels 0; Gaps 0;

Qy 1414 gcgtccagctggtgcagagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1473
Db 1347 gagaccagcctgggctgctgcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1406

Qy 1474 aagtggtcaactccactgctgtgtcccgaggagcagcagcagcagcagcagcagcagcagcag 1533
Db 1407 tctacagggcgctgccactgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1456

Qy 1534 cggctgcgtgtgcagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1593
Db 1467 gagaagcgtctctggtggaatagagcagcagcagcagcagcagcagcagcagcagcagcagcag 1526

Qy 1594 ctcttctgcgtctctctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 1653
Db 1527 ctcttctgcgtctctctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 1586

Qy 1654 gactaccagatgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1713
Db 1587 cagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1646

Qy 1714 ctggccaacttttccaag 1731
Db 1647 attgggaacctgggccaag 1664

```

```

RESULT 4
US-09-050-863-2
; Sequence 2, Application US/09050863
; Patent No. 6114111
; GENERAL INFORMATION:
; APPLICANT: Lao, Ying
; APPLICANT: Hiang, Betty
; APPLICANT: Payan, Don
; TITLE OF INVENTION: Mammalian Protein Interaction Cloning
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hohnbach, Test, Albritton & Herbert

```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 18, 2001, 07:58:32 ; Search time 6516.34 Seconds
(without alignments)
4445.587 Million cell updates/sec

Title: US-09-294-298-3
Perfect score: 4134
Sequence: 1 taagggccccccacccgacc.....tgggggtgctatccccatcct 4134

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 7991742 seqs, 3503743858 residues

Total number of hits satisfying chosen parameters: 15983484

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:
1: gb_estl1:
2: gb_est2:
3: gb_est3:
4: gb_est4:
5: gb_est5:
6: gb_est6:
7: gb_est7:
8: gb_est8:
9: gb_est9:
10: gb_est10:
11: gb_est11:
12: gb_est12:
13: gb_est13:
14: gb_est14:
15: gb_est15:
16: gb_est16:
17: gb_est17:
18: gb_est18:
19: gb_est19:
20: gb_est20:
21: gb_est21:
22: gb_est22:
23: gb_est23:
24: gb_est24:
25: gb_est25:
26: gb_est26:
27: gb_est27:
28: gb_est28:
29: gb_est29:
30: gb_est30:
31: gb_est31:
32: gb_est32:
33: gb_est33:
34: gb_est34:
35: gb_est35:
36: gb_est36:
37: gb_est37:
38: gb_est38:
39: gb_est39:
40: gb_est40:
41: en_estba:
42: en_estfun:
43: en_esthum1:
44: en_esthum2:
45: en_esthum3:
46: en_esthum4:
47: en_esthum5:
48: en_esthum6:
49: en_esthum7:
50: en_esthum8:
51: en_esthum9:
52: en_esthum10:
53: en_esthum11:
54: en_esthum12:
55: en_esthum13:
56: en_esthum14:
57: en_esthum15:
58: en_esthum16:
59: en_esthum17:
60: en_esthum18:
61: en_esthum19:
62: en_esthum20:
63: en_estin1:
64: en_estin2:
65: en_estin3:
66: en_estin4:
67: en_estov1:
68: en_estov2:
69: en_estpl1:
70: en_estpl2:
71: en_estpl3:
72: en_estpl4:
73: en_estpl5:
74: en_estroi:
75: en_estro2:
76: en_estro3:
77: en_estro4:
78: en_estro5:
79: en_estro6:
80: en_estro7:
81: en_estro8:
82: en_estro9:
83: en_estroi0:
84: en_estroi1:
85: en_estroi2:
86: en_estroi3:
87: gb_est41:
88: gb_est42:
89: gb_est43:
90: gb_est44:
91: gb_est45:
92: gb_est46:
93: gb_est47:
94: gb_est48:
95: gb_est49:
96: gb_est50:
97: gb_est51:
98: gb_est52:
99: gb_est53:
100: gb_est54:
101: gb_est55:
102: gb_est56:
103: gb_est57:
104: gb_est67:
105: gb_est68:
106: gb_est69:
107: gb_est70:
108: gb_est71:
109: gb_est72:
110: gb_est73:
111: gb_est74:
112: en_esthum21:
113: en_esthum22:
114: en_esthum23:
115: en_estom1:
116: en_estom2:

117: em_estpl6:*
 118: em_estpl7:*
 119: em_estpl8:*
 120: em_estpl9:*
 121: em_estpl10:*
 122: em_estpl11:*
 123: em_estpl12:*
 124: em_estpl13:*
 125: em_estpl14:*
 126: em_estpl15:*
 127: em_estpl16:*
 128: em_estpl17:*
 129: em_estpl18:*
 130: em_estpl19:*
 131: em_estpl20:*
 132: em_estpl21:*
 133: em_estpl22:*
 134: em_estpl23:*
 135: em_estpl24:*
 136: em_estpl25:*
 137: em_estpl26:*
 138: em_estpl27:*
 139: em_estpl28:*
 140: em_estpl29:*
 141: em_estpl30:*
 142: em_estpl31:*
 143: em_estpl32:*
 144: em_estpl33:*
 145: em_estpl34:*
 146: em_estpl35:*
 147: em_estpl36:*
 148: em_estpl37:*
 149: em_estpl38:*
 150: em_estpl39:*
 151: em_estpl40:*
 152: em_estpl41:*
 153: em_estpl42:*
 154: em_estpl43:*
 155: em_estpl44:*
 156: em_estpl45:*
 157: em_estpl46:*
 158: em_estpl47:*
 159: em_estpl48:*
 160: em_estpl49:*
 161: em_estpl50:*
 162: em_estpl51:*
 163: em_estpl52:*
 164: em_estpl53:*
 165: em_estpl54:*
 166: em_estpl55:*
 167: em_estpl56:*
 168: em_estpl57:*
 169: em_estpl58:*
 170: em_estpl59:*
 171: em_estpl60:*
 172: em_estpl61:*
 173: em_estpl62:*
 174: em_estpl63:*
 175: em_estpl64:*
 176: em_estpl65:*
 177: em_estpl66:*
 178: em_estpl67:*
 179: em_estpl68:*
 180: em_estpl69:*
 181: em_estpl70:*
 182: em_estpl71:*
 183: em_estpl72:*
 184: em_estpl73:*
 185: em_estpl74:*
 186: em_estpl75:*
 187: em_estpl76:*
 188: em_estpl77:*
 189: em_estpl78:*

190: gb_gss25:*
 191: gb_gss26:*
 192: gb_gss27:*
 193: gb_gss28:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	490.4	11.9	571	37	AV590698	AV590698 AV590698
2	383.8	9.3	474	181	AZ399131	AZ399131 IM0164P14
3	294.6	7.1	593	95	AW779747	AW779747 hm85f10.x
4	250.4	6.1	463	108	BE463433	BE463433 hw23d06.x
5	244.6	5.9	427	11	AA709010	AA709010 zf94h11.s
6	233.6	5.7	597	90	AW421227	AW421227 f194f10.x
7	228.8	5.5	682	173	AQ993964	AQ993964 RPCI-23-2
8	222	5.4	566	22	AI606711	AI606711 ml58d01.y
9	221.8	5.4	413	10	AA691827	AA691827 vt05d01.r
10	214	5.2	707	28	AJ396547	AJ396547 AJ396547
11	211.8	5.1	555	27	AI958829	AI958829 fg22d01.y
12	201	4.9	628	108	BE490923	BE490923 db38g08.x
13	197.2	4.8	516	111	BE695645	BE695645 MRI-BT080
14	195.8	4.7	347	87	AW204612	AW204612 UI-H-B11-
15	195.4	4.7	596	136	BE848541	BE848541 uw39f05.y
16	190	4.6	481	135	BE773015	BE773015 RCI-FT013
17	185	4.5	542	93	AW656799	AW656799 109245 MA
18	182.2	4.4	296	23	AI650331	AI650331 wa18f01.x
19	177.5	4.3	345	39	AW076911	AW076911 f103d08.y
20	173	4.2	511	22	AI577567	AI577567 UI-R-Y0-V
21	167.2	4.0	388	183	BE68206	BE68206 CIT9785K-A-
22	165.6	4.0	474	87	AW205989	AW205989 UI-H-B11-
23	161.2	3.9	460	22	AI609604	AI609604 tw91e09.x
24	157.8	3.8	1020	191	CNS020VGW	AL215753 Tetraodon
25	157	3.8	449	9	AA554056	AA554056 n101g07.s
26	154.6	3.7	440	2	AA107246	AA107246 ml58d01.r
27	149.4	3.6	651	19	AI327335	AI327335 mp74c11.x
28	148.6	3.6	877	192	CNS04D10	AL385885 Tetraodon
29	146.8	3.6	443	40	AW136165	AW136165 UI-H-B11-
30	146.8	3.6	464	95	AW828497	AW828497 fk62h05.x
31	146	3.5	452	9	AA554055	AA554055 n101g06.s
32	145.4	3.5	407	133	BE015978	BE015978 fk62h05.y
33	139.8	3.4	382	22	AI562171	AI562171 vw73f04.x
34	138.6	3.4	317	91	AW527364	AW527364 UI-R-B01-
35	138.6	3.4	457	2	AA118397	AA118397 mp74c11.r
36	138.4	3.3	268	92	AW593427	AW593427 bq15e10.x
37	136.6	3.3	379	9	AA575948	AA575948 nm56e05.s
38	135.6	3.3	554	111	BE723198	BE723198 192821 MA
39	133.4	3.2	336	26	AI674961	AI674961 ul27f03.x
40	131.8	3.2	694	28	AJ399084	AJ399084 AJ399084
41	130.4	3.2	335	134	BE085114	BE085114 CM2-BT066
42	130	3.1	1006	190	CNS00HWF	AL073826 Drosophila
43	128.2	3.1	631	191	CNS01X05	AL172022 Tetraodon
44	126.8	3.1	975	191	CNS03ANH	AL235430 Tetraodon
45	124.2	3.0	397	111	BE695651	BE695651 MRI-BT080

ALIGNMENTS

RESULT 1
 AV590698
 LOCUS AV590698 Bos taurus brain fetus 571 bp mRNA
 DEFINITION AV590698 Bos taurus brain fetus 5' mRNA sequence.
 ACCSSION AV590698
 VERSION AV590698.1 GI:9701691
 KEYWORDS EST.
 SOURCE COW.

29-AUG-2000
 EST
 EMBL014B04

ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bos.

REFERENCE 1 (bases 1 to 571)
AUTHORS Sugimoto, Y., Hirotsune, S., Takasuga, A., Itoh, R., Jitohzono, A. and Suzuki, H.

TITLE Bovine cDNA sequencing
JOURNAL Unpublished (2000)
COMMENT Contact: Yoshikazu Sugimoto
Animal Genetics Division
Shirakawa Institute of Animal Genetics

Odakura, Nishigo, Nishi-shirakawa, Fukushima 961-8061, Japan
Tel: 81-248-25-5641
Fax: 81-248-25-5725

Email: kazusugie@ocn.ne.jp
Single pass sequencing.

This clone was obtained from a polyA-deleted cDNA library.

FEATURES
source

1..571

Location/Qualifiers

/organism="Bos taurus"

/db_xref="taxon:9913"

/clone="E1BR014B04"

/clone_lib="Bos taurus brain fetus"

/cissue_type="brain"

/dev_stage="fetus"

/lab_host="DH10B"

/note="Vector: pZ11; Site_1: SalI; Site_2: NotI; Poly A

was deleted from a NotI site"

161 c 169 g 116 t 1 others

BASE COUNT 124 a

ORIGIN

Query Match 11.9%; Score 490.4; DB 37; Length 571;
Best Local Similarity 93.2%; Pred. No. 2e-103;
Matches 523; Conservative 0; Mismatches 37; Indels 1; Gaps 1;

Qy 1145 qcaaggaggggtcgtagtgcactggttcacatctctgcaaaacacagcaggaagcgaag 1204

Db 11 GCACAGGAGGTCGCCATGCTGCTGGTTCACATCTTCAGAGTACAGGACGCAAGG 70

Qy 1205 actctcttcagacatggccatgttcagaggttagaccggttcatgtgagcggaacacctca 1264

Db 71 ACTTCTTTTCAGACATGGCCATGTCGAGGTGGACCGGTTTCATGGAACGAGACACCTCA 130

Qy 1265 tattccgcgagacacgcctcgccactaaagccatagagagatgatagactatggcc 1324

Db 131 TATTCCGCGAGACACGCTCGCCACTAAGCCATAGAGAGTACATGAGACTGATTGGTC 190

Qy 1325 agaatactcagagtgccattggggaggttcacccggctctgtgtatgaatctgaggaga 1384

Db 191 AGAAATACCTCAAGGATGCCATCGGGGAATTCATCCGCTGCTCTGTATGAATCTGAGGAGA 250

Qy 1385 actgtaagttagaccacatcaagtgcacagcgtccagctctgagagacacagcgaacc 1444

Db 251 ACTGTGAGTGGACCCCATCAAGTCGACGGCATCCAGTTGGCTGACACACGACCAACC 310

Qy 1445 tgcgagtgtctgtgagttggccctgtgcaaggttggttaactccatctggtttcccca 1504

Db 311 TGGCGATGTGCTGTGAGTTGGCCCTGTGCAAGGTGGTCAACTCCCATTTGCGTGTCCCGA 370

Qy 1505 gggagctgaaggaggtgtttgcatcatggtcggtgtgctgtgagagcggtggtgaggg 1564

Db 371 GGGAGCTGAAGGAGGTGTTTGGCTCTTGGCAGCTGCGCTGTGACAGCGGGCGGAGG 430

Qy 1565 acattgctgacaggtgtatcagcgtctgctcttctctgcttctctgcttctgctgcccgaatca 1624

Db 431 ACATTGGGACCGGCTGATGAGCGGCTGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 490

Qy 1625 tgtgcgccactgtgtt-tggactgtgagaggtaccagatgacagagcagcagcagcagcagc 1683

Db 491 TGTGCGCCACGCTTTTCGGGGCTCATGCGAGGTACCCAGATGAGCAAACTTCGAGAAC 550

Qy 1684 ctacccctcatcgccaaggtt 1704

Db 551 CTCACGCTCATCGCCAAGGTT 571

RESULT 2

AZ399131/c

LOCUS

DEFINITION

1M0164P14R Mouse 10kb plasmid UUGC1M library Mus musculus genomic

clone UUGC1M0164P14 R, DNA sequence.

ACCESSION

AZ399131

VERSION

AZ399131.1

KEYWORDS

GSS

SOURCE

house mouse.

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 474)

AUTHORS

Duni, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,

Islam, H., Longacre, S., Mahmood, M., Meenen, E., Pedersen, T., Reilly

, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.

and Wright, D., Weiss, R.

Mouse whole genome scaffolding with paired end reads from 10kb

plasmid inserts

JOURNAL

Unpublished (2000)

COMMENT

Contact: Robert B. Weiss

University of Utah Genome Center

University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT

84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0164 row: P column: 14

Seq primer: CACACGAGAAACAGCTATGACC

Class: plasmid ends

High quality sequence stop: 474.

Location/Qualifiers

1..474

/organism="Mus musculus"

/strain="C57BL/6j"

/db_xref="taxon:10090"

/clone="UUGC1M0164P14"

/clone_lib="Mouse 10kb plasmid UUGC1M library"

/sex="Male"

/lab_host="E. Coli strain XL10-Gold, TI-resistant, P-

/note="Vector: PWD42nv; Purified genomic DNA from M.

musculus C57BL/6j (male) was obtained from the Jackson

Laboratory Mouse DNA Resource

(http://www.jax.org/resources/documents/dnares/). The DNA

was hydrodynamically sheared by repeated passage through a

0.005 inch orifice at constant velocity. The sheared DNA

was blunt end-repaired with T4 DNA polymerase and T4

polynucleotide kinase. Adaptor oligonucleotides were

ligated to the blunt ends in high molar excess. The

adaptor DNA was purified and size-selected for a 9.5 to

10.5 kb range using preparative agarose gel

electrophoresis. Vector DNA was prepared from a derivative

of PWD42 (gi14732114|gb|AF129072.1), a copy-number

inducible derivative of plasmid R1. The vector was ligated

with adaptors complementary to the insert adaptors and

purified. The sheared, adaptor mouse DNA was annealed to

adaptor vector DNA, and transformed into

chemically-competent E. coli XL10-Gold (Stratagene) cells

and selected for ampicillin resistance."

BASE COUNT 92 a 151 c 131 g 100 t

ORIGIN

Query Match 9.3%; Score 383.8; DB 181; Length 474;

Best Local Similarity 93.7%; Pred. No. 1.1e-78;

Matches 400; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

•

•

•

•

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 18, 2001, 19:58:55 ; Search time 505.11 Seconds
(without alignments)
3074.557 Million cell updates/sec

Title: US-09-294-298-3
Perfect score: 4134
Sequence: 1 taagggcccccacccgacc.....tggggtgctatccccatcct 4134

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 480022 seqs, 187831343 residues
Total number of hits satisfying chosen parameters: 960044

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_36.*
1: /SID56/gcgdata/geneseq/geneseq/NA1980.DAT.*
2: /SID56/gcgdata/geneseq/geneseq/NA1981.DAT.*
3: /SID56/gcgdata/geneseq/geneseq/NA1982.DAT.*
4: /SID56/gcgdata/geneseq/geneseq/NA1983.DAT.*
5: /SID56/gcgdata/geneseq/geneseq/NA1984.DAT.*
6: /SID56/gcgdata/geneseq/geneseq/NA1985.DAT.*
7: /SID56/gcgdata/geneseq/geneseq/NA1986.DAT.*
8: /SID56/gcgdata/geneseq/geneseq/NA1987.DAT.*
9: /SID56/gcgdata/geneseq/geneseq/NA1988.DAT.*
10: /SID56/gcgdata/geneseq/geneseq/NA1989.DAT.*
11: /SID56/gcgdata/geneseq/geneseq/NA1990.DAT.*
12: /SID56/gcgdata/geneseq/geneseq/NA1991.DAT.*
13: /SID56/gcgdata/geneseq/geneseq/NA1992.DAT.*
14: /SID56/gcgdata/geneseq/geneseq/NA1993.DAT.*
15: /SID56/gcgdata/geneseq/geneseq/NA1994.DAT.*
16: /SID56/gcgdata/geneseq/geneseq/NA1995.DAT.*
17: /SID56/gcgdata/geneseq/geneseq/NA1996.DAT.*
18: /SID56/gcgdata/geneseq/geneseq/NA1997.DAT.*
19: /SID56/gcgdata/geneseq/geneseq/NA1998.DAT.*
20: /SID56/gcgdata/geneseq/geneseq/NA1999.DAT.*
21: /SID56/gcgdata/geneseq/geneseq/NA2000.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Query match length DB ID Description

Result No.	Score	Match	Length	DB ID	Description
1	95.8	2.3	2647	18 T76739	CDNA encoding acti
2	63.6	2.3	1337	20 Z17263	Human gene express
3	61	1.5	795	19 V55830	FLGA insert stabl
4	61	1.5	799	19 V55831	Nucleotide sequenc
5	61	1.5	5452	20 X0923	Anti-sense strand
6	61	1.5	8705	20 Z23778	Vector pShuttle DN
7	61	1.5	9600	19 V21683	Vector plasmid pCM
8	61	1.5	10380	20 Z22248	Nucleotide sequenc
9	61	1.5	10596	14 Q51731	Plasmid pCisEBON f
10	61	1.5	10596	17 T40348	Plasmid pCisEBON f
11	61	1.5	10596	20 X15650	Nucleotide sequenc
12	59.6	1.4	1925	20 X90924	Epstein Barr Virus

13	59.4	1.4	1908	8 N71064	Gene encoding Plas
14	59.2	1.4	1000	21 A02484	Human colon cancer
15	57.8	1.4	4999	20 Z23895	Murine LOBO homolo
16	57.6	1.4	2338	19 V23249	Nephila clavipes s
17	57.6	1.4	2338	21 Z38195	N. clavipes spider
18	56	1.4	2000	8 N71065	Gene encoding Plas
19	56	1.4	2338	12 Q14183	N. clavipes draglin
20	56	1.4	18442	18 X83006	Partial mouse WRN
21	55.8	1.3	2004	18 T83356	Nephila clavipes s
22	54.6	1.3	32207	20 V73805	KSHV LOR DNA (nucl
23	54.6	1.3	137507	19 V19941	KSHV long unique c
24	54.4	1.3	9551	20 Z22301	CDNA encoding a hu
25	54	1.3	543	13 Q23092	Antigen tc-7a gene
26	53.6	1.3	1995	12 Q14184	N. clavipes draglin
27	53.6	1.3	1995	19 V23250	N. clavipes spider
28	53.6	1.3	1995	21 Z38196	N. clavipes spider
29	53.4	1.3	4055	20 Z40484	Human ZC2 DNA. HO
30	53.4	1.3	4180	21 A10669	Human protein kina
31	53	1.3	1137	14 Q33061	Plasmodium vivax c
32	52.6	1.3	10732	21 A10594	Gene encoding a su
33	52.4	1.3	114955	20 X53491	Human adenosine A1
34	52.2	1.3	1847	19 V04694	Cell cycle protein
35	51.4	1.2	3198	20 X02974	Human IL-1ra BAC c
36	51.2	1.2	2744	16 Q98470	MisP1-containing p
37	51.2	1.2	51259	18 X83007	Partial mouse WRN
38	50.2	1.2	1686	16 Q87587	DNA encoding Leuco
39	49	1.2	2040	18 T62137	Leishmania brazili
40	49	1.2	2040	19 V47557	Leishmania antigen
41	48.8	1.2	1218	21 A02488	Human colon cancer
42	48.8	1.2	2313	20 Z07197	Human lung tumour
43	48.6	1.2	1432	19 V47585	Leishmania antigen
44	47.8	1.2	3337	17 T34620	P. vivax ESP-1 blo
45	47.8	1.2	3337	20 X15174	DNA encoding a sec

ALIGNMENTS

RESULT 1
T76739
ID T76739 standard; cDNA; 2647 BP.
XX
AC T76739;
DT 08-OCT-1997 (first entry)
DE CDNA encoding active type R-Ras protein binding protein p98.
XX
XX Active-type-R-Ras protein; p98; tumour; disease; cancer;
KW cellular Ca ion; phospholipid; tumour forming promoter;
KW tumour forming inhibitor; ss.
XX Bos taurus.
XX
XX Key Location/Qualifiers
XX CDS 13..2517
XX FT /*tag= a
XX FT /product= p98
XX
XX JF09135688-A.
XX
XX PD 27-MAY-1997.
XX
XX PF 16-NOV-1995; 95JP-0298720.
XX
XX PR 16-NOV-1995; 95JP-0298720.
XX
XX PA (KIRI) KIRIN BREWERY KK.
XX
XX DR WPI: 1997-335993/31.
XX DR P-PSDB; W24227.
XX
XX PT Active R-Ras protein-binding protein p98 - used in the treatment of

